

151590

**Schreiber, David**

---

**From:** Ramirez, Delia  
**Sent:** Friday, April 15, 2005 6:51 PM  
**To:** Schreiber, David  
**Subject:** case 10/652334

Hi,

I would like to request the following search (commercial and interference): SEQ ID NO:1-9 in the protein databases.

Thank you,

---

Delia M. Ramirez, Ph.D.  
Patent Examiner  
Recombinant Enzymes-Art Unit 1652  
USPTO  
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70  
Alexandria, VA 22314  
(571) 272-0938  
delia.ramirez@uspto.gov

**This Page Blank (uspto)**

FOR OFFICIAL USE ONLY

Scientific and Technical Information Center  
**SEARCH REQUEST FORM**

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
Art Unit: \_\_\_\_\_ Phone Number: 2- \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Location (Bldg/Room#): \_\_\_\_\_ (Mailbox #): \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK  
\*\*\*\*\*

To ensure an efficient and quality search, please attach a copy of the cover sheet, claims, and abstract or fill out the following:

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Date: \_\_\_\_\_

**Search Topic:**

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known.

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

\*\*\*\*\*

**STAFF USE ONLY**

	Type of Search	Vendors and cost where applicable
Searcher: <u>D. Schreiber</u>	_____ NA Sequence (#)	_____ STN _____ Dialog
Searcher Phone #: <u>272-2524</u>	<u>9</u> AA Sequence (#)	_____ Questel/Orbit _____ Lexis/Nexis
Searcher Location: <u>Remsen E01 A4</u>	_____ Structure (#)	_____ Westlaw _____ WWW/Internet
Date Searcher Picked Up: _____	_____ Bibliographic	<input checked="" type="checkbox"/> In-house sequence systems <u>See page</u>
Date Completed: <u>4/25</u>	_____ Litigation	<input checked="" type="checkbox"/> Commercial _____ Oligomer _____ Score/Length
Searcher Prep & Review Time: <u>15</u>	_____ Fulltext	<input checked="" type="checkbox"/> Interference _____ SPDI _____ Encode/Transl
Online Time: <u>6</u>	_____ Other	_____ Other (specify)

*This Page Blank (aspto)*



CC antiviral therapy and for suppression of pathological nonsense mutations.

CC The present sequence is *Saccharomyces cerevisiae* motif peptide

XX Sequence 9 AA; 95.7%; Score 45; DB 7; Length 9;  
 SQ Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Query Match  
 1 GPPGRTKXT 9  
 1 GPPGRTKXT 9  
 Db 1 GPPGRTKXT 9

RESULT 2  
 ADP44103  
 ID ADP44103 standard; peptide; 9 AA.  
 AC ADP44103;  
 XX  
 DT 18-NOV-2004 (first entry)  
 DE Yeast translation termination modulation protein motif I.  
 XX gene therapy; translation termination; RNA helicase; MTT1;  
 KM frameshift frequency; aberrant transcript degradation;  
 KM peptidyl transferase modulation; beta-thalassemia; beta-globin;  
 KM Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
 KM Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;  
 KM Ovarian Cancer; Wilms Tumour; Hirschprung disease; Cystic fibrosis;  
 KM Kidney Stone; Familial hypercholesterolemia; Retinitis Pigmentosa;  
 KM Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; Yeast.  
 OS Saccharomycetes cerevisiae.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 8 /note= "Any amino acid"  
 FT  
 XX US2004115787-A1.  
 PD 17-JUN-2004.  
 XX  
 PF 28-AUG-2003; 2003US-00652334.  
 XX  
 PR 22-JUL-1998; 98US-0093685P.  
 PR 22-JUL-1999; 99US-00359266.  
 XX  
 PA (PELTZ/) PELTZ S.  
 PA (CZAP/) CZAPLINSKI K.  
 PA (DIMM/) DIMMAN J D.  
 XX  
 PI Pelcz S, Czaplinski K, Dimman JD;  
 DR WPI; 2004-449400/42.  
 XX  
 PT Identifying a test composition or agent that modulates the efficiency of  
 PT translation termination comprises contacting the MTT1 with the test  
 PT composition or agent, and determining if the test composition or agent  
 PT inhibits the MTT1.  
 XX  
 PS Claim 33; SEQ ID NO 1; 41pp; English.  
 XX  
 CC The invention relates to a method of identifying a test composition that  
 CC modulates the efficiency of translation termination comprising contacting  
 CC the RNA helicase MTT1 with a composition or agent under conditions  
 CC permitting binding between the MTT1 and the composition, detecting  
 CC specific binding of the test composition or agent to the MTT1, and  
 CC determining if the test composition or agent inhibits the MTT1. The  
 CC composition and methods are useful for modulating the fidelity of  
 CC translation termination or for identifying agents that affect the  
 CC functional activity of mRNAs by altering frameshift frequency, permit  
 CC monitoring of a termination event, promote degradation of aberrant  
 CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl

CC transferase activity during initiation, elongation, termination and mRNA  
 CC degradation of translation. The agents, which may be antagonists or  
 CC agonists, are useful in screening, diagnostic and therapeutic purposes,  
 CC for diseases or conditions resulting from or cause premature translation,  
 CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular  
 CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,  
 CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,  
 CC Hirschprung disease, Cystic fibrosis, Kidney Stones, Familial  
 CC hypercholesterolemia, Retinitis Pigmentosa, or Neurofibromatosis,  
 CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents  
 CC the amino acid sequence of the yeast translation termination modulation  
 CC protein motif I.  
 XX  
 SQ Sequence 9 AA; 95.7%; Score 45; DB 8; Length 9;  
 Query Match  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Query Match  
 1 GPPGRTKXT 9  
 1 GPPGRTKXT 9  
 Db 1 GPPGRTKXT 9

RESULT 3  
 AAY77804  
 ID AAY77804 standard; peptide; 10 AA.  
 AC AAY77804;  
 XX  
 DT 31-MAY-2000 (first entry)  
 DE Motif I comprised in a gene modulating translation termination.  
 XX  
 FH Helicase B; HCSB; MTT1; modulator of translation termination; eRF1, eRF3;  
 KM eukaryotic release factor; peptidyl transferase; beta-thalassemia;  
 KM beta-globin; Duchenne/Becker Muscular Dystrophy; anislemic.  
 XX  
 OS Unidentified.  
 OS  
 FH Key Location/Qualifiers  
 FT Misc-difference 1.10 /note= "residues indicated Xaa are unspecified"  
 FT  
 XX WO200005586-A2.  
 PD 03-FEB-2000.  
 XX  
 PF 22-JUL-1999; 99WO-US016802.  
 XX  
 PR 22-JUL-1998; 98US-00120435.  
 XX  
 PA (UNIV-) UNIV NEW JERSEY.  
 XX  
 PI Pelcz S, Czaplinski K, Dimman JD;  
 DR WPI; 2000-171458/15.  
 XX  
 PT New multiprotein complex which can modulate peptidyl transferase activity  
 PT during translation, useful to treat diseases associated with peptidyl  
 PT transferase activity e.g. Duchenne/Becker Muscular Dystrophy.  
 XX  
 PS Claim 33; Page 79; 89pp; English.  
 XX  
 CC The invention provides a new multiprotein complex which can modulate  
 CC peptidyl transferase activity during translation. The complex comprises  
 CC the gene encoding Helicase B (HCSB; renamed MTT1, for Modulator of  
 CC Translation Termination) and the conserved proteins known to interact and  
 CC carry out translation termination in eukaryotic cells, peptidyl  
 CC eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to  
 CC modulate peptidyl transferase activity during translation in a cell. It  
 CC can be administered therapeutically combined with a carrier in  
 CC pharmaceutical compositions to treat diseases associated with peptidyl

CC transferase activity, especially diseases resulting from a nonsense or  
CC frameshift mutation e.g. beta-thalassemia, beta-globin. Duchene/Becker  
CC Muscular Dystrophy etc. It can be used to identify disease conditions  
CC involving a defect in the complex, by transfecting cells with encoding  
CC nucleic acid and determining the proportion of defective complex before  
CC and after transfection. It is also useful to screen for drugs involved in  
CC peptidyl transferase activity during translation, inhibiting the  
CC interaction between MT11 and eRF3 or involved in enhancing translation  
CC termination. Vectors comprising polynucleotides encoding the complex (or  
CC antisense sequences) can be constructed and introduced into cells to  
CC interfere with complex expression and so modulate the efficiency of  
CC translation termination of mRNA and/or degradation of aberrant  
CC transcripts in a cell. Agents binding to the complex can be identified  
CC and included in therapeutic compositions useful as above, and/or used to  
CC modulate peptidyl transferase activity during translation in cells. They  
CC are also useful to modulate the efficiency of translation termination of  
CC mRNA at a nonsense codon and/or promote degradation of aberrant  
CC transcripts in cells. The method can be used to identify agents/  
CC compositions modulating binding to MT11, useful to identify genes.  
CC Sequences AAY7804-812 represent motifs 1-IX comprised in the genes of  
CC interest, used for modulating translation termination

CC Sequence 10 AA;

Query Match 95.7%; Score 45; DB 3; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGKTKYT 9  
|||  
Db 1 GPGGKTKYT 9

RESULT 4

ADFS9025  
ID ADFS9025 standard; protein; 159 AA.

AC ADFS9025;

DT 12-FEB-2004 (first entry)

XX Human polypeptide sequence SEQ ID NO:1433.

XX biological activity; genetic engineering; hybridisation probe; oligomer;  
KM primer; chromosome mapping; gene mapping; recombinant protein production;  
XX human.

OS Homo sapiens.

PN WO2003080795-A2.

XX 02-OCT-2003.

PP 09-AUG-2002; 2002WO-US025485.

PR 09-AUG-2001; 2001US-0311261P.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Yang Y, Wang Z, Wang G, Ma Y;

DR WPI; 2003-876918/81.

DR N-PSDB; ADFS8025.

XX New polynucleotides, useful as hybridization probes, oligomers or  
PT primers, for chromosome or gene mapping, for the recombinant production  
PT of proteins, and for generating antisense DNA or RNA.

XX Claim 20; SEQ ID NO 1433; 571pp; English.

CC The present sequence represents a polypeptide (II) with biological  
CC activity, which is encoded by an isolated polynucleotide sequence (I)  
CC from the present invention. Also described: (I) a vector comprising (I);

CC (2) an expression vector comprising (I); (3) a host cell genetically  
CC engineered to comprise (I) which is operatively associated with a  
CC regulatory sequence that modulates expression of (I) in the host cell;  
CC (4) a polypeptide (II) encoded by (I); (5) a composition comprising the  
CC polypeptide of (4) and a carrier; (6) an antibody directed against the  
CC polypeptide of (4); (7) detecting (I) or the polypeptide of (4) in a  
CC sample; (8) identifying a compound that binds to the polypeptide of (4);  
CC (9) producing the polypeptide of (4); and (10) a collection of  
CC polynucleotides comprising at least one of the polynucleotide sequences  
CC (I). The polynucleotides (I) can be used as hybridisation probes,  
CC oligomers or primers, for chromosome or gene mapping, for the recombinant  
CC production of proteins, and for generating antisense DNA or RNA.

CC Sequence 159 AA;

Query Match 83.0%; Score 39; DB 7; Length 159;  
Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPGGKTKYT 9  
|||  
Db 140 GPGGKTKYT 148

RESULT 5

ABP38224  
ID ABP38224 standard; protein; 135 AA.

AC ABP38224;

DT 24-JUL-2002 (first entry)

XX Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3069.

XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
KM antibacterial; gene therapy.

OS Staphylococcus epidermidis.

PN US6380370-B1.

XX 30-APR-2002.

PP 13-AUG-1998; 98US-00134001.

PR 14-AUG-1997; 97US-0055779P.

PR 08-NOV-1997; 97US-0064964P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Doucette-Stamm LA, Bush D;

DR WPI; 2002-381255/41.

DR N-PSDB; ABN90769.

XX Novel isolated nucleic acid encoding a Staphylococcus epidermis  
PT polypeptide, useful for diagnosing and treating bacterial infections.

XX Disclosure; SEQ ID NO 3069; 267pp; English.

XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences  
CC given in ABP3124 to ABP37960. The S. epidermidis sequences have  
CC antibacterial activity and can be used in gene therapy. The sequences can  
CC also be used in the diagnosis and treatment of bacterial infections,  
CC particularly S. epidermidis infections. The sequences can be used to  
CC screen for compounds able to interfere with the S. epidermidis life cycle  
CC or inhibit S. epidermidis infection. N.B. The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from the USPTO web site

XX Sequence 135 AA;

CC

Query Match 78.7%; Score 37; DB 5; Length 135;  
 Best Local Similarity 77.8%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GPGTKTXYT 9  
 |||||  
 DB 83 GPGTKTIT 91

RESULT 6  
 AAG82625  
 ID AAG82625 standard; protein; 278 AA.

AC AAG82625;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:2344.

KW Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;  
 endocarditis.

OS Staphylococcus epidermidis.

PN W0200134809-A2.

PD 17-MAY-2001.

PF 09-NOV-2000; 2000WO-US030782.

PR 09-NOV-1999; 99US-0164258P.

PA (GLAXO ) GLAXO GROUP LTD.

PI kimmerly WJ;

DR WPI; 2001-316495/33.

DR N-PSDB; AAH53475.

PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
 useful for vaccinating against infections, e.g. endocarditis.

PS Claim 18; Page 620; 2186pp; English.

CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
 (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)

CC and (II) can have antibacterial activity and therefore can be used in  
 vaccination. The nucleic acids (II) may be used to produce the S.

CC epidermidis polypeptides (II) via the production of vectors containing  
 them which are used to produce host cells which express the

CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
 used to vaccinate subjects and to raise antibodies against the bacteria.

CC The polypeptides may also be used to assay for other inhibitors of their  
 activity and therefore identify compounds that may be used for the

CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to  
 AAH5090 represent specifically claimed S. epidermidis genomic DNA

CC polynucleotide sequences from the present invention. AAH5091 to AAH5098  
 represent oligonucleotide sequences and primers which are used in the

CC exemplification of the present invention. N.B. The present invention  
 specifically claims all the polynucleotide sequences given in the

CC sequence listing of the present specification, however the sequence  
 listing only goes up to SEQ ID NO:4454 so even though sequences are given

CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present  
 for SEQ ID NO:4455 to 4464

CC SQ Sequence 278 AA;

Query Match 78.7%; Score 37; DB 4; Length 278;  
 Best Local Similarity 77.8%; Pred. No. 4.4e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GPGTKTXYT 9  
 |||||

DB 44 GPGTKTIT 52

RESULT 7  
 ABP38223  
 ID ABP38223 standard; protein; 383 AA.

AC ABP38223;

DT 24-JUL-2002 (first entry)

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3068.

KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
 antibacterial; gene therapy.

OS Staphylococcus epidermidis.

PN US6380370-B1.

PD 30-APR-2002.

PF 13-AUG-1998; 98US-00134001.

PR 14-AUG-1997; 97US-0055779P.

PR 08-NOV-1997; 97US-0064964P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Doucette-Stamm LA, Bush D;

DR WPI; 2002-381255/41.

DR N-PSDB; ABN90768.

PT Novel isolated nucleic acid encoding a Staphylococcus epidermis  
 polypeptide, useful for diagnosing and treating bacterial infections.

PS Disclosure; SEQ ID NO 3068; 267pp; English.

CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
 frame (ORF) nucleic acid sequences which encode the amino acid sequences

CC given in ABP3124 to ABP3760. The S. epidermidis sequences have  
 CC antibacterial activity and can be used in gene therapy. The sequences can

CC also be used in the diagnosis and treatment of bacterial infections,  
 CC particularly S. epidermidis infections. The sequences can be used to

CC screen for compounds able to interfere with the S. epidermidis life cycle  
 CC or inhibit S. epidermidis infection. N.B. The sequence data for this

CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from the USPTO web site

SQ Sequence 383 AA;

Query Match 78.7%; Score 37; DB 5; Length 383;  
 Best Local Similarity 77.8%; Pred. No. 6e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GPGTKTXYT 9  
 |||||  
 DB 83 GPGTKTIT 91

RESULT 8  
 AAG82772  
 ID AAG82772 standard; protein; 404 AA.

AC AAG82772;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:2638.

KW Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;  
 endocarditis.

XX	Staphylococcus epidermidis.
OS	
XX	WO200134809-A2.
PN	
XX	17-MAY-2001.
PD	
XX	09-NOV-2000; 2000WO-US030762.
PF	
XX	09-NOV-1999; 99US-0164258P.
PR	
PA	(GLAX ) GLAXO GROUP LTD.
PI	Kimmerly MJ;
DR	WPI; 2001-316495/33.
DR	N-P5DB; AAH53622.
XX	
PT	Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT	useful for vaccinating against infections, e.g. endocarditis.
XX	
PS	Claim 18; Page 692; 2188pp; English.
XX	
CC	AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC	(II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
CC	and (II) can have antibacterial activity and therefore can be used in
CC	vaccination. The nucleic acids (I) may be used to produce the S.
CC	epidermidis polypeptides (II) via the production of vectors containing
CC	them which are used to produce host cells which express the
CC	polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC	used to vaccinate subjects and to raise antibodies against the bacteria.
CC	The polypeptides may also be used to assay for other inhibitors of their
CC	activity and therefore identify compounds that may be used for the
CC	treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC	AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC	polynucleotide sequences from the present invention. AAH5091 to AAH5098
CC	represent oligonucleotide sequences and primers which are used in the
CC	exemplification of the present invention. N.B. The present invention
CC	specifically claims all the polynucleotide sequences given in the
CC	sequence listing of the present specification, however the sequence
CC	listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC	in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC	for SEQ ID NO:4455 to 4464
XX	
SO	Sequence 404 AA;
XX	
Query Match	78.7%; Score 37; DB 4; Length 404;
Best local Similarity	77.8%; Pred.No. 6.3e+02;
Matches 7; Conservative	0; Mismatches 2; Indels 0; Gaps 0.
OY	1 GPGTKTXY 9           44 GEPGKTTT 52
Dd	
XX	
RESULT 9	
ID	AAG82701 standard; protein; 470 AA.
AC	AAG82701;
DT	03-SEP-2001 (first entry)
DE	S. epidermidis open reading frame protein sequence SEQ ID NO:2496.
KM	Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination; endocarditis.
XX	
OS	Staphylococcus epidermidis.
NN	WO200134809-A2.
XX	
DD	17-MAY-2001.

XX	09-NOV-2000; 2000WO-US030782.
PF	
XX	
PR	09-NOV-1999; 99US-0164258P.
PA	(GLAX ) GLAXO GROUP LTD.
XX	
PI	Kimmerly MJ;
XX	
DR	WPI, 2001-316495/33.
XX	N-PsDB; AAH53551.
PT	Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT	useful for vaccinating against infections, e.g. endocarditis.
PS	Claim 18; Page 657; 2188pp; English.
XX	
AAH52304	to AAH51970 represent nucleic acid (I) encoding polypeptides
CC	(II), given in AAG81454 to AAG81120, from Staphylococcus epidermidis. (I)
CC	and (II) can have antibacterial activity and therefore can be used in
CC	vaccination. The nucleic acid (I) may be used to produce the S.
CC	epidermidis polypeptides (II) via the production of vectors containing
CC	them which are used to produce host cells which express the
CC	polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC	used to vaccinate subjects and to raise antibodies against the bacteria.
CC	The polypeptides may also be used to assay for other inhibitors of their
CC	activity and therefore identify compounds that may be used for the
CC	treatment of S. epidermidis infections, e.g. endocarditis. AAH51971 to
CC	AAH5090 represent specifically claimed S. epidermidis genomic DNA
CC	polynucleotide sequences from the present invention. AAH5091 to AAH5098
CC	represent oligonucleotide sequences and primers which are used in the
CC	exemplification of the present invention. N.B. The present invention
CC	specifically claims all the polynucleotide sequences given in the
CC	sequence listing of the present specification, however the sequence
CC	listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC	in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC	for SEQ ID NO:4455 to 4464
SQ	Sequence 470 AA;
XX	
Query Match	78.7%; Score 37; DB 4; Length 470;
Best Local Similarity	77.8%; Pred. No. 7.3e+02;
Matches	7; Conservative 0; Mismatches 2; Indels 0; Gaps 0.
QY	1 GPGTKTXYT 9           
Db	44 GPGTKTIT 52
RESULT 10	
ID	ADG30695 standard; protein; 577 AA.
AC	ADG30695;
XX	
DT	26-FEB-2004 (first entry)
DE	Xanthomonas axonopodis pv citri plant pathology-related XAC2118 protein.
XX	
KM	Xanthomonas microorganism; plant; pathology; bacterial pest; Xac; Xcc; XAc.
OS	Xanthomonas axonopodis pv. citri.
PN	MO2003089647-A1.
PD	30-OCT-2003.
PF	22-APR-2003; 2003WO-BR000060.
PR	22-APR-2002; 2002US-0374620P.
PA	(AMPA-) FUNDAÇÃO AMPARO À PESQUISA DO ESTADO.

XX Da Silva ACR, Parah SC, Quaggio RB, Reinach FDC, Ferro JA;  
PI De Oliveira JCF, De Lata ML, Setubal JC, Furlan LR;  
XX WPI; 2003-865444/80.  
DR N-PSDB; ADG30694.  
XX  
PT New nucleic acid molecule from a Xanthomonas microorganism, useful in  
determining the presence of Xanthomonas bacteria in a sample.  
PS Claim 8; SEQ ID NO 52; 145bp; English.  
XX  
XX The invention relates to a novel isolated nucleic acid molecule from a  
Xanthomonas microorganism where the nucleic acid molecule is associated  
with pathogenicity caused by the Xanthomonas microorganism, or its  
variant, that causes reduced or enhanced pathogenicity. The nucleic acid  
of the invention may be useful in detecting the presence of Xanthomonas  
bacteria in a sample, as well as in plant pathology, for example, for  
identifying nucleic acid molecules and proteins involved in pathology  
caused by bacterial pests. The current sequence is that of the  
Xanthomonas axonopodis pv. citri (Xac) plant pathology-related XAC  
protein of the invention.  
XX  
SQ Sequence 577 AA;  
Query Match 78.7%; Score 37; DB 7; Length 577;  
Best Local Similarity 66.7%; Pred. No. 8.8e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
OY 1 GPGGTXXT 9  
DB 460 GPGGSETDT 468  
RESULT 11  
AAG82914  
ID AAG82914 standard; protein; 696 AA.  
XX  
XX AAG82914;  
XX  
DT 03-SEP-2001 (first entry)  
XX  
DE S. epidermidis open reading frame protein sequence SEQ ID NO:2922.  
XX  
KW Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;  
KM endocarditis.  
XX  
OS Staphylococcus epidermidis.  
XX  
PN WO200134809-A2.  
XX  
PD 17-MAY-2001.  
XX  
PE 09-NOV-2000; 2000WO-US030782.  
XX  
PR 09-NOV-1999; 99US-0164258P.  
XX  
PA (GLAX ) GLAXO GROUP LTD.  
XX  
PI Kimmery WJ;  
XX  
PS WPI; 2001-316495/33.  
DR N-PSDB; AAH53764.  
XX  
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
useful for vaccinating against infections, e.g. endocarditis.  
XX  
PS Claim 18; Page 763-764; 2188bp; English.  
XX  
XX AAH52304 to AAH5970 represent nucleic acids (I) encoding polypeptides  
(II), given in AAG81454 to AAG8120, from Staphylococcus epidermidis. (I)  
and (II) can have antibacterial activity and therefore can be used in  
vaccination. The nucleic acids (I) may be used to produce the S.

CC epidermidis polypeptides (II) via the production of vectors containing  
them which are used to produce hosts cells which express the  
polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
used to vaccinate subjects and to raise antibodies against the bacteria.  
The polypeptides may also be used to assay for other inhibitors of their  
activity and therefore identify compounds that may be used for the  
treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to  
AAH5090 represent specifically claimed S. epidermidis genomic DNA  
polynucleotide sequences from the present invention. AAH5091 to AAH5098  
represent oligonucleotide sequences and primers which are used in the  
amplification of the present invention. N.B. The present invention  
specifically claims all the polynucleotide sequences given in the  
sequence listing of the present specification, however the sequence  
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given  
in the disclosure for SEQ ID NO:4455 to 4472, no sequences are present  
for SEQ ID NO:4455 to 4464  
XX  
SQ Sequence 696 AA;  
Query Match 78.7%; Score 37; DB 4; Length 696;  
Best Local Similarity 77.8%; Pred. No. 1e+03;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 GPGGTXXT 9  
DB 94 GPGGTXITT 102  
RESULT 12  
AAG81738  
ID AAG81738 standard; protein; 715 AA.  
XX  
XX AAG81738;  
XX  
DT 03-SEP-2001 (first entry)  
XX  
DE S. epidermidis open reading frame protein sequence SEQ ID NO:570.  
XX  
KW Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;  
KM endocarditis.  
XX  
OS Staphylococcus epidermidis.  
XX  
PN WO200134809-A2.  
XX  
PD 17-MAY-2001.  
XX  
PE 09-NOV-2000; 2000WO-US030782.  
XX  
PR 09-NOV-1999; 99US-0164258P.  
XX  
PA (GLAX ) GLAXO GROUP LTD.  
XX  
PI Kimmery WJ;  
XX  
PS WPI; 2001-316495/33.  
DR N-PSDB; AAH52588.  
XX  
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
useful for vaccinating against infections, e.g. endocarditis.  
XX  
PS Claim 18; Page 189; 2188bp; English.  
XX  
XX AAH52304 to AAH5970 represent nucleic acids (I) encoding polypeptides  
(II), given in AAG81454 to AAG8120, from Staphylococcus epidermidis. (I)  
and (II) can have antibacterial activity and therefore can be used in  
vaccination. The nucleic acids (I) may be used to produce the S.  
epidermidis polypeptides (II) via the production of vectors containing  
them which are used to produce hosts cells which express the  
polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
used to vaccinate subjects and to raise antibodies against the bacteria.  
The polypeptides may also be used to assay for other inhibitors of their  
activity and therefore identify compounds that may be used for the

CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to  
 CC AAH5090 represent specifically claimed *S. epidermidis* genomic DNA  
 CC polynucleotide sequences from the present invention. AAH5091 to AAH5098  
 CC represent oligonucleotide sequences and primers which are used in the  
 CC exemplification of the present invention. N.B. The present invention  
 CC specifically claims all the polynucleotide sequences given in the  
 CC sequence listing of the present specification, however the sequence  
 CC listing only goes up to SEQ ID NO:4454 so even though sequences are given  
 CC in the disclosure for SEQ ID NO:4455 to 4472, no sequences are present  
 CC for SEQ ID NO:4455 to 4464

XX SQ Sequence 715 AA;

Query Match 78.7%; Score 37; DB 4; Length 715;  
 Best Local Similarity 77.8%; Pred. No. 1.1e+03;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPGGKTKYT 9  
 Db 44 GPGGKTKYT 52

RESULT 13  
 ID ABP38222 standard; protein; 846 AA.

XX AC ABP38222;

XX DT 24-JUL-2002 (first entry)

XX DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3067.

XX KM Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;

XX KW antibacterial; gene therapy.

XX OS Staphylococcus epidermidis.

XX FN US6380370-B1.

XX PD 30-APR-2002.

XX PP 13-AUG-1998; 98US-00134001.

XX PR 14-AUG-1997; 97US-0055779P.

XX PR 08-NOV-1997; 97US-0064964P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Doucette-Stamm LA, Bush D;

XX DR WPI; 2002-381255/41.

XX DR N-PSDB; ABN90767.

PT Novel isolated nucleic acid encoding a *Staphylococcus epidermidis*  
 PT polypeptide, useful for diagnosing and treating bacterial infections.

PS Disclosure; SEQ ID NO 3067; 267pp; English.

XX ABN90538 to ABN93374 represent *Staphylococcus epidermidis* open reading  
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences  
 CC given in ABP5124 to ABP37960. The *S. epidermidis* sequences have  
 CC antibacterial activity and can be used in gene therapy. The sequences can  
 CC also be used in the diagnosis and treatment of bacterial infections,  
 CC particularly *S. epidermidis* infections. The sequences can be used to  
 CC screen for compounds able to interfere with the *S. epidermidis* life cycle  
 CC or inhibit *S. epidermidis* infection. N.B. The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from the USPTO web site

XX SQ Sequence 846 AA;

Query Match 78.7%; Score 37; DB 5; Length 846;  
 Best Local Similarity 77.8%; Pred. No. 1.3e+03;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 GPGGKTKYT 9  
 Db 34 GPGGKTKYT 42

RESULT 14  
 ID AAG83007 standard; protein; 910 AA.

XX AC AAG83007;

XX DT 03-SEP-2001 (first entry)

XX DE *S. epidermidis* open reading frame protein sequence SEQ ID NO:3108.

XX KM Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;  
 KW endocarditis.

XX OS Staphylococcus epidermidis.

XX PN WO200134809-A2.

XX PD 17-MAY-2001.

XX PP 09-NOV-2000; 2000WO-US030782.

XX PR 09-NOV-1999; 99US-0164258P.

XX PA (GLAX ) GLAXO GROUP LTD.

XX PI Kimerly WJ;

XX DR WPI; 2001-316495/33.

XX DR N-PSDB; AAH53857.

PT Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*,  
 PT useful for vaccinating against infections, e.g. endocarditis.

PS Claim 18; Page 819; 218pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
 CC (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*. (I)  
 CC and (II) can have antibacterial activity and therefore can be used in  
 CC vaccination. The nucleic acids (I) may be used to produce the *S.*  
 CC *epidermidis* polypeptides (II) via the production of vectors containing  
 CC them which are used to produce hosts cells which express the  
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
 CC used to vaccinate subjects and to raise antibodies against the bacteria.  
 CC The polypeptides may also be used to assay for other inhibitors of their  
 CC activity and therefore identify compounds that may be used for the  
 CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to  
 CC AAH5090 represent specifically claimed *S. epidermidis* genomic DNA  
 CC polynucleotide sequences from the present invention. AAH5091 to AAH5098  
 CC represent oligonucleotide sequences and primers which are used in the  
 CC exemplification of the present invention. N.B. The present invention  
 CC specifically claims all the polynucleotide sequences given in the  
 CC sequence listing of the present specification, however the sequence  
 CC listing only goes up to SEQ ID NO:4454 so even though sequences are given  
 CC in the disclosure for SEQ ID NO:4455 to 4472, no sequences are present  
 CC for SEQ ID NO:4455 to 4464

XX SQ Sequence 910 AA;

Query Match 78.7%; Score 37; DB 4; Length 910;  
 Best Local Similarity 77.8%; Pred. No. 1.3e+03;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPGGKTKYT 9  
 Db 94 GPGGKTKYT 102

## RESULT 15

AD505211 standard; protein, 997 AA.

AD505211;

04-NOV-2004 (first entry)

Staphylococcus epidermis polypeptide seqid 4506.

antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis; recombinant expression vector; infection; computer readable medium; computer based system.

Staphylococcus epidermidis.

US2004147734-A1.

29-JUL-2004.

01-DEC-2003; 2003US-00724972.

08-NOV-1997; 97US-0064964P.

13-AUG-1998; 98US-00134001.

29-NOV-1999; 99US-00450969.

(DOUC/) DOUCETTE-STMM L.

(BUSH/) BUSH D.

Doucette-Stamm L, Bush D;

WPI; 2004-580138/56.

N-PSDB; ADS01439.

New isolated polypeptide and encoding nucleic acid derived from

Staphylococcus epidermidis, useful for diagnosing, preventing and/or

treating an S. epidermidis bacterial infection.

Claim 17; SEQ ID NO 4506; 741bp; English.

The invention describes an isolated nucleic acid comprising a nucleotide sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO: 1-3772) and encoding an Staphylococcus epidermidis polypeptide with any of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as given in the specification. Also described are: a recombinant expression vector; a cell comprising a recombinant expression vector of (1); producing an S. epidermidis polypeptide; an isolated nucleic acid comprising a nucleotide sequence of at least 8 nucleotides in length; a vaccine composition for prevention or treatment of an S. epidermidis infection, comprising a nucleic acid cited above and a carrier; treating a subject for S. epidermidis infection; a recombinant or substantially pure preparation of an S. epidermidis polypeptide or its fragment; a vaccine composition for prevention or treatment of an S. epidermidis infection; detecting the presence of a Staphylococcus nucleic acid in a sample; a computer readable medium having recorded in it the nucleotide sequences with SEQ ID NO: 1-3772 or its fragments; a computer based system for identifying fragments of the Staphylococcus genome of commercial importance; a computer based system for identifying fragments of the Staphylococcus plasmids of commercial importance; identifying commercially important nucleic acid fragments of the Staphylococcus genome and/or plasmids; and identifying an expression modulating fragment of the Staphylococcus genome and/or plasmids. The methods and compositions of the present invention are useful for the diagnosis, prevention and/or treatment of an Staphylococcal epidermidis bacterial infection. This is the amino acid sequence of a S. epidermis protein of the invention.

Sequence 997 AA;

Query Match 78.7%; Score 37; DB 8; Length 997;  
Best Local Similarity 77.8%; Pred. No. 1.5e+03;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GPGTGTXT 9  
| | | | |  
Db 697 GPGTGTIT 705

Search completed: April 18, 2005, 08:03:40  
Job time : 46.914 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 07:37:21 ; Search time 11.236 Seconds  
(without alignments)  
59.807 Million cell updates/sec

Title: US-10-652-334-1

Perfect score: 47

Sequence: 1 GPGTKXTX 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:  
1: /cgn2\_6/ptocata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptocata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptocata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptocata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptocata/1/1aa/PTCUS.COMB.pep:\*  
6: /cgn2\_6/ptocata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	95.7	9	4	US-09-359-268A-1
2	37	78.7	135	3	US-09-134-001C-3069
3	37	78.7	278	4	US-09-710-279-2344
4	37	78.7	383	3	US-09-134-001C-3068
5	37	78.7	404	4	US-09-710-279-2638
6	37	78.7	470	4	US-09-710-279-2496
7	37	78.7	696	4	US-09-710-279-2322
8	37	78.7	715	4	US-09-710-279-570
9	37	78.7	846	3	US-09-134-001C-3067
10	37	78.7	910	4	US-09-710-279-3108
11	37	78.7	1183	3	US-09-134-001C-3530
12	35.5	75.5	340	4	US-09-538-092-1103
13	35.5	75.5	363	4	US-09-538-092-1072
14	35.5	75.5	415	4	US-09-359-268A-25
15	35.5	75.5	917	4	US-09-248-796A-19347
16	35.5	75.5	993	4	US-09-538-092-1100
17	35	74.5	116	4	US-09-902-540-10762
18	35	74.5	173	4	US-09-902-540-1840
19	35	74.5	420	4	US-09-252-991A-19117
20	35	74.5	793	4	US-09-270-767-42801
21	34.5	73.4	28	2	US-08-724-354D-6
22	34.5	73.4	28	2	US-09-359-268A-6
23	34.5	73.4	380	4	US-09-359-268A-29
24	34.5	73.4	414	4	US-09-359-268A-28
25	34.5	73.4	472	4	US-09-359-268A-26
26	34.5	73.4	683	4	US-09-538-092-483
27	34.5	73.4	971	2	US-08-724-354D-22

28	34.5	73.4	971	3	US-09-270-984A-22	Sequence 22, Appl
29	34.5	73.4	971	3	US-09-177-431-8	Sequence 8, Appl
30	34.5	73.4	1043	2	US-08-724-354D-4	Sequence 4, Appl
31	34.5	73.4	1043	3	US-09-270-984A-4	Sequence 2, Appl
32	34.5	73.4	1118	2	US-08-724-354D-2	Sequence 2, Appl
33	34.5	73.4	1118	2	US-09-270-984A-2	Sequence 2, Appl
34	34.5	73.4	1140	4	US-09-949-016-10116	Sequence 10116, A
35	34	72.3	156	4	US-09-270-767-49434	Sequence 49434, A
36	34	72.3	156	4	US-09-270-767-49434	Sequence 49434, A
37	34	72.3	231	4	US-09-902-540-12103	Sequence 12103, A
38	34	72.3	348	4	US-09-902-540-14234	Sequence 14234, A
39	34	72.3	802	4	US-09-252-991A-25050	Sequence 25050, A
40	33	70.2	60	3	US-09-314-268-143	Sequence 143, App
41	33	70.2	62	3	US-09-268-364-6	Sequence 6, Appl
42	33	70.2	65	3	US-09-268-364-6	Sequence 25779, A
43	33	70.2	139	4	US-09-252-991A-25779	Sequence 58172, A
44	33	70.2	163	4	US-09-270-767-58172	Sequence 4, Appl
45	33	70.2	213	3	US-09-268-364-4	

## ALIGNMENTS

RESULT 1  
US-09-359-268A-1  
; Sequence 1, Application US/09359268A  
; Patent No. 6630294  
; GENERAL INFORMATION:  
; APPLICANT: Peltz, Stuart  
; APPLICANT: Czaplinski, Kevin  
; APPLICANT: Dimman, Jonathan D.  
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
; FILE REFERENCE: 601-1-85N  
; CURRENT APPLICATION NUMBER: US/09/359,268A  
; CURRENT FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 60/093,685  
; PRIOR FILING DATE: 1998-07-22  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
; FEATURES:  
; OTHER INFORMATION: Xaa = any amino acid  
US-09-359-268A-1  
Query Match 95.7%; Score 45; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cq 1 GPGTKXTX 9  
DB 1 GPGTKXTX 9  
RESULT 2  
US-09-134-001C-3069  
; Sequence 3069, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 3069  
LENGTH: 135  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3069

Query Match 78.7%; Score 37; DB 3; Length 135;  
Best Local Similarity 77.8%; Pred. No. 54;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GPGTKTXXT 9  
| | | | |  
DB 83 GPGTKTIT 91

RESULT 3  
US-09-710-279-2344  
Sequence 2344, Application US/09710279  
Patent No. 6703492  
GENERAL INFORMATION:  
APPLICANT: KIMBERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: P03480US  
CURRENT APPLICATION NUMBER: US/09/710,279  
CURRENT FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: 60/164,258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 2344  
LENGTH: 278  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-710-279-2344

Query Match 78.7%; Score 37; DB 4; Length 278;  
Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GPGTKTXXT 9  
| | | | |  
DB 44 GPGTKTIT 52

RESULT 4  
US-09-134-001C-3068  
Sequence 3068, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 3068  
LENGTH: 383  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3068

Query Match 78.7%; Score 37; DB 3; Length 383;  
Best Local Similarity 77.8%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GPGTKTXXT 9  
| | | | |  
DB 83 GPGTKTIT 91

RESULT 5  
US-09-710-279-2638  
Sequence 2638, Application US/09710279  
Patent No. 6703492  
GENERAL INFORMATION:  
APPLICANT: KIMBERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: P03480US  
CURRENT APPLICATION NUMBER: US/09/710,279  
CURRENT FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: 60/164,258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 2638  
LENGTH: 404  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-710-279-2638

Query Match 78.7%; Score 37; DB 4; Length 404;  
Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GPGTKTXXT 9  
| | | | |  
DB 44 GPGTKTIT 52

RESULT 6  
US-09-710-279-2496  
Sequence 2496, Application US/09710279  
Patent No. 6703492  
GENERAL INFORMATION:  
APPLICANT: KIMBERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: P03480US  
CURRENT APPLICATION NUMBER: US/09/710,279  
CURRENT FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: 60/164,258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 2496  
LENGTH: 470  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: amino acid sequence  
FEATURE:  
NAME/KEY: MOD RES  
LOCATION: (470)  
OTHER INFORMATION: variable amino acid  
US-09-710-279-2496

Query Match 78.7%; Score 37; DB 4; Length 470;  
Best Local Similarity 77.8%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GPGTKTXXT 9  
| | | | |  
DB 44 GPGTKTIT 52

RESULT 7  
US-09-710-279-2922  
; Sequence 2922, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PUS480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2922  
; LENGTH: 696  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: amino acid sequence  
US-09-710-279-2922

Query Match 78.7%; Score 37; DB 4; Length 696;  
Best Local Similarity 77.8%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPGTKTXXT 9  
Db 94 GPGTKTIT 102

RESULT 8  
US-09-710-279-570  
; Sequence 570, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PUS480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 570  
; LENGTH: 715  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: amino acid sequence  
US-09-710-279-570

Query Match 78.7%; Score 37; DB 4; Length 715;  
Best Local Similarity 77.8%; Pred. No. 2.7e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPGTKTXXT 9  
Db 44 GPGTKTIT 52

RESULT 9  
US-09-134-001C-3067  
; Sequence 3067, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 3067  
; LENGTH: 846  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3067

Query Match 78.7%; Score 37; DB 3; Length 846;  
Best Local Similarity 77.8%; Pred. No. 3.2e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPGTKTXXT 9  
Db 34 GPGTKTIT 42

RESULT 10  
US-09-710-279-3108  
; Sequence 3108, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PUS480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3108  
; LENGTH: 910  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: amino acid sequence  
US-09-710-279-3108

Query Match 78.7%; Score 37; DB 4; Length 910;  
Best Local Similarity 77.8%; Pred. No. 3.4e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPGTKTXXT 9  
Db 94 GPGTKTIT 102

RESULT 11  
US-09-134-001C-3530  
; Sequence 3530, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 3530  
; LENGTH: 1183

TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3530

Query Match 76.7%; Score 37; DB 3; Length 1183;  
Best Local Similarity 77.8%; Pred. No. 4.4e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GPGT-KTXXT 9  
Db 697 GPGTKTST 705

RESULT 12  
US-09-538-092-1103  
Sequence 1103, Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:  
APPLICANT: Gluc, Lolc  
APPLICANT: Mansfield, Traci A.  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: Curataseqformatter Version 0.9  
SEQ ID NO 1103  
LENGTH: 340  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (0)...(0)  
OTHER INFORMATION: Polypeptide Accession Number P40937  
US-09-538-092-1103

Query Match 75.5%; Score 35.5; DB 4; Length 340;  
Best Local Similarity 80.0%; Pred. No. 2.3e+02;  
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GPGT-KTXXT 9  
Db 60 GPGTKTST 69

RESULT 13  
US-09-538-092-1072  
Sequence 1072, Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:  
APPLICANT: Gluc, Lolc  
APPLICANT: Mansfield, Traci A.  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: Curataseqformatter Version 0.9  
SEQ ID NO 1072  
LENGTH: 363  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (0)...(0)

OTHER INFORMATION: Polypeptide Accession Number P35249  
US-09-538-092-1072

Query Match 75.5%; Score 35.5; DB 4; Length 363;  
Best Local Similarity 80.0%; Pred. No. 2.5e+02;  
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GPGT-KTXXT 9  
Db 78 GPGTKTST 87

RESULT 14  
US-09-359-268A-25  
Sequence 25, Application US/09359268A  
Patent No. 6630294  
GENERAL INFORMATION:  
APPLICANT: Peltz, Stuart  
APPLICANT: Czaplinski, Kevin  
APPLICANT: Dimman, Jonathan D.  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/09/359,268A  
CURRENT FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
PRIOR FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 25  
LENGTH: 415  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-359-268A-25

Query Match 75.5%; Score 35.5; DB 4; Length 415;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GPGT-KTXXT 9  
Db 3 GPGTKTST 12

RESULT 15  
US-09-248-796A-19347  
Sequence 19347, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Kelch Weinstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 19347  
LENGTH: 917  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-19347

Query Match 75.5%; Score 35.5; DB 4; Length 917;  
Best Local Similarity 80.0%; Pred. No. 6e+02;  
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GPGT-KTXXT 9  
Db 3 GPGTKTST 12

Tue Apr 19 09:10:38 2005

us-10-652-334-1.ra1

Page 5

Db 473 GPCGNGKTST 482

Search completed: April 18, 2005, 08:18:31  
Job time : 12.2336 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 08:06:16 ; Search time 31.927 Seconds  
(without alignments)  
93.693 Million cell updates/sec

Title: US-10-652-334-1  
Perfect score: 47  
Sequence: 1 GPGTKXTX 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues  
Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCTI\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10D\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	95.7	9	US-10-652-334-1	Sequence 1, Appl1
2	37	78.7	126	US-10-424-599-144310	Sequence 143107, A
3	37	78.7	344	US-10-424-599-176807	Sequence 176807, A
4	37	78.7	434	US-10-767-701-44355	Sequence 44355, A
5	37	78.7	563	US-10-437-963-156740	Sequence 156740, A
6	37	78.7	577	US-10-418-8618-52	Sequence 52, Appl1
7	37	78.7	593	US-10-437-963-107474	Sequence 107474, A
8	37	78.7	1030	US-10-470-0488-296	Sequence 296, Appl1
9	37	78.7	1245	US-10-172-502-19	Sequence 19, Appl1
10	37	78.7	2397	US-10-282-122A-71232	Sequence 71232, A
11	36	76.6	140	US-10-425-114-57982	Sequence 57982, A
12	36	76.6	832	US-10-408-765A-1776	Sequence 1776, Ap
13	36	76.6	1330	US-10-108-260A-3237	Sequence 3237, Ap

14	36	76.6	1626	US-10-202-167-2	Sequence 2, Appl1
15	36	76.6	4455	US-10-287-226-304	Sequence 304, Ap
16	35.5	75.5	139	US-10-767-701-36594	Sequence 36594, A
17	35.5	75.5	188	US-10-767-701-41880	Sequence 41880, A
18	35.5	75.5	224	US-10-264-049-2354	Sequence 2354, Ap
19	35.5	75.5	256	US-10-424-599-145467	Sequence 145467, A
20	35.5	75.5	300	US-10-767-701-38318	Sequence 38318, A
21	35.5	75.5	311	US-10-369-493-12479	Sequence 12479, A
22	35.5	75.5	319	US-10-369-493-21404	Sequence 21404, A
23	35.5	75.5	336	US-10-424-599-210517	Sequence 210517, A
24	35.5	75.5	339	US-10-437-963-176776	Sequence 176776, A
25	35.5	75.5	340	US-10-369-493-1941	Sequence 1941, Ap
26	35.5	75.5	344	US-10-437-963-109094	Sequence 109094, A
27	35.5	75.5	353	US-10-369-493-22192	Sequence 22192, A
28	35.5	75.5	363	US-10-369-493-3826	Sequence 3826, Ap
29	35.5	75.5	367	US-10-425-114-64115	Sequence 64115, A
30	35.5	75.5	369	US-10-369-493-18667	Sequence 18667, A
31	35.5	75.5	371	US-10-369-493-3730	Sequence 3730, Ap
32	35.5	75.5	382	US-10-437-963-109098	Sequence 109098, A
33	35.5	75.5	383	US-10-425-114-42743	Sequence 42743, A
34	35.5	75.5	394	US-10-369-493-13298	Sequence 13298, A
35	35.5	75.5	409	US-10-369-493-955	Sequence 955, Ap
36	35.5	75.5	415	US-10-652-334-25	Sequence 25, Appl1
37	35.5	75.5	447	US-10-369-493-5730	Sequence 5790, Ap
38	35.5	75.5	611	US-10-369-493-10288	Sequence 10288, A
39	35.5	75.5	626	US-10-437-963-166322	Sequence 166322, A
40	35.5	75.5	637	US-10-424-599-233501	Sequence 233501, A
41	35.5	75.5	639	US-10-425-114-37717	Sequence 37717, A
42	35.5	75.5	642	US-10-425-114-37557	Sequence 37557, A
43	35.5	75.5	781	US-10-437-963-176775	Sequence 176775, A
44	35	74.5	93	US-10-437-963-180141	Sequence 180141, A
45	35	74.5	111	US-10-180-410-10	Sequence 10, Appl1

#### ALIGNMENTS

RESULT 1  
US-10-652-334-1  
Sequence 1, Application US/10652334  
Publication No. US20040115787A1  
GENERAL INFORMATION:  
APPLICANT: Peltz, Stuart  
APPLICANT: Czaplinski, Kevin  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/10/652,334  
CURRENT FILING DATE: 2003-08-28  
PRIOR APPLICATION NUMBER: US/09/359,268A  
PRIOR FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
FEATURE:  
OTHER INFORMATION: Xaa = any amino acid  
US-10-652-334-1

Query Match 95.7%; Score 45; DB 16; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0;

QY 1 GPGTKXTX 9  
DB 1 GPGTKXTX 9

RESULT 2  
US-10-424-599-144310  
; Sequence 144310, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 144310  
; LENGTH: 126  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(126)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_101325C.1.pep  
US-10-424-599-144310

Query Match 78.7%; Score 37; DB 15; Length 126;  
Best Local Similarity 85.7%; Pred. No. 1.6e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGTKT 7  
Db 79 GPGTKT 85

RESULT 3  
US-10-424-599-176807  
; Sequence 176807, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 176807  
; LENGTH: 344  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_130674C.1.pep  
US-10-424-599-176807

Query Match 78.7%; Score 37; DB 15; Length 344;  
Best Local Similarity 77.8%; Pred. No. 4.3e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPGTKT 9  
Db 8 GPGTKT 16

RESULT 4  
US-10-767-701-44355  
; Sequence 44355, Application US/10767701  
; Publication No. US20040172684A1

; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 44355  
; LENGTH: 434  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C2240\_1.pep  
US-10-767-701-44355

Query Match 78.7%; Score 37; DB 16; Length 434;  
Best Local Similarity 85.7%; Pred. No. 5.4e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGTKT 7  
Db 261 GPGSKT 267

RESULT 5  
US-10-437-963-156740  
; Sequence 156740, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 156740  
; LENGTH: 563  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_56378C.1.pep  
US-10-437-963-156740

Query Match 78.7%; Score 37; DB 16; Length 563;  
Best Local Similarity 75.0%; Pred. No. 6.9e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPGTKT 9  
Db 12 PPGTKT 19

RESULT 6  
US-10-418-861B-52  
; Sequence 52, Application US/10418861B  
; Publication No. US20040010131A1  
; GENERAL INFORMATION:  
; APPLICANT: da Silva, Ana Claudia Rasera  
; APPLICANT: Farah, Shaker Chuck  
; APPLICANT: Quaggio, Ronaldo Bento  
; APPLICANT: Reinach, Fernando de Castro  
; APPLICANT: Ferro, Jesus Aparecido



```

; APPLICANT: De Oliveira, Julio Cezar Franco
; APPLICANT: De Laia, Marcelo Luiz
; APPLICANT: Setubal Jose C.
; APPLICANT: Purián, Luiz Roberto
; TITLE OF INVENTION: Isolated Xanthomonas nucleic acid molecules, proteins encoded the
; FILE REFERENCE: PAPER 205.1 US
; CURRENT APPLICATION NUMBER: US/10/418,861B
; CURRENT FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: US 60/374,620
; PRIOR FILING DATE: 2002-04-22
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 52
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Xanthomonas
; FEATURE:
US-10-418-861B-52
```

```

Query Match      78.7%; Score 37; DB 15; Length 577;
Best Local Similarity 66.7%; Pred. No. 7.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 GPGTKT 9
        |||||
Db      460 GPGSETDT 468
```

```

RESULT 7
US-10-437-963-107474
; Sequence 107474, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Mu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 107474
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_11822C.1.Dep
US-10-437-963-107474
```

```

Query Match      78.7%; Score 37; DB 16; Length 593;
Best Local Similarity 85.7%; Pred. No. 7.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GPGTKT 7
        |||||
Db      420 GPGSKT 426
```

```

RESULT 8
US-10-470-048B-296
; Sequence 296, Application US/10470048B
; Publication No. US20050037444A1
; GENERAL INFORMATION:
; APPLICANT: MEINKE ET AL.
; TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF
; TITLE OF INVENTION: ANTIGENS TO A SPECIFIC PATHOGEN
; FILE REFERENCE: SONN:035US
```

```

; CURRENT APPLICATION NUMBER: US/10/470, 048B
; CURRENT FILING DATE: 2003-07-25
; NUMBER OF SEQ ID NOS: 603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 296
; LENGTH: 1030
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (265)..(271)
; OTHER INFORMATION: X = anything
US-10-470-048B-296
```

```

Query Match      78.7%; Score 37; DB 17; Length 1030;
Best Local Similarity 77.8%; Pred. No. 1.3e+03;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 GPGTKT 9
        |||||
Db      44 GPGTKTIT 52
```

```

RESULT 9
US-10-172-502-19
; Sequence 19, Application US/10172502
; Publication No. US20030185833A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy et al.
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES. . .
; FILE REFERENCE: P07263US01/BAS
; CURRENT APPLICATION NUMBER: US/10/172,502
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/298,038
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 1245
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-172-502-19
```

```

Query Match      78.7%; Score 37; DB 14; Length 1245;
Best Local Similarity 77.8%; Pred. No. 1.5e+03;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 GPGTKT 9
        |||||
Db      771 GPGTKTIT 779
```

```

RESULT 10
US-10-282-122A-71232
; Sequence 71232, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyckind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
```

PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: Patent version 3.1  
SEQ ID NO 71232  
LENGTH: 2397  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-10-282-122A-71232

Query Match 78.7%; Score 37; DB 15; Length 2397;  
Best Local Similarity 77.8%; Pred. No. 2.9e+03;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GPGTKTXYT 9  
DB 771 GPGTKTXYT 779

RESULT 11  
US-10-425-114-57982  
Sequence 57982, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jindong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313)B  
CURRENT FILING DATE: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 57982  
LENGTH: 140  
TYPE: PRT  
ORGANISM: Zea mays subsp. mexicana  
FEATURE:  
OTHER INFORMATION: Clone ID: UC-ZMROTESINT024D01\_FLI.pep  
US-10-425-114-57982

Query Match 76.6%; Score 36; DB 15; Length 140;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GPGTK 6  
DB 41 GPGTK 46

RESULT 12

US-10-408-765A-1776  
Sequence 1776, Application US/10408765A  
Publication No. US20040101874A1  
GENERAL INFORMATION:  
APPLICANT: Ghosh, Soumitra S.  
APPLICANT: Fahy, Colin D.  
APPLICANT: Zhang, Bing  
APPLICANT: Gibson, Bradford W.  
APPLICANT: Taylor, Steven W.  
APPLICANT: Glenn, Gary M.  
APPLICANT: Warnock, Dale E.  
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
FILE REFERENCE: 660088.465  
CURRENT APPLICATION NUMBER: US/10/408,765A  
CURRENT FILING DATE: 2003-04-04  
NUMBER OF SEQ ID NOS: 3077  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1776  
LENGTH: 832  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-408-765A-1776

Query Match 76.6%; Score 36; DB 16; Length 832;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GPGTK 6  
DB 92 GPGTK 97

RESULT 13  
US-10-108-260A-3237  
Sequence 3237, Application US/10108260A  
Publication No. US20040005560A1  
GENERAL INFORMATION:  
APPLICANT: HELIX RESEARCH INSTITUTE  
TITLE OF INVENTION: No. US20040005560A1 full length cDNA  
FILE REFERENCE: H1-A0106  
CURRENT APPLICATION NUMBER: US/10/108,260A  
CURRENT FILING DATE: 2002-03-27  
NUMBER OF SEQ ID NOS: 5458  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 3237  
LENGTH: 1330  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-108-260A-3237

Query Match 76.6%; Score 36; DB 15; Length 1330;  
Best Local Similarity 100.0%; Pred. No. 2.4e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GPGTK 6  
DB 1046 GPGTK 1051

RESULT 14  
US-10-202-167-2  
Sequence 2, Application US/10202167  
Publication No. US20030143564A1  
GENERAL INFORMATION:  
APPLICANT: Burgess, Robert Eugene  
APPLICANT: Koch, Manuel  
APPLICANT: Bruckner-Tuderman, Leena  
APPLICANT: Keene, Douglas R.  
APPLICANT: Brunken, William Joseph  
TITLE OF INVENTION: COLLAGEN XXII, A NOVEL HUMAN COLLAGEN  
TITLE OF INVENTION: AND USES THEREOF  
FILE REFERENCE: 10287-072001

```

; CURRENT APPLICATION NUMBER: US/10/202,167
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/309,158
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-202-167-2

```

```

Query Match          76.6%; Score 36; DB 14; Length 1626;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GPGTK 6
        |||||
Db      762 GPGTK 767

```

```

RESULT 15
US-10-287-226-304
; Sequence 304, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
; APPLICANT: Agee, Michele L.,
; APPLICANT: Alsobrook, John P.,
; APPLICANT: Berghs, Constance,
; APPLICANT: Boldog, Ference,
; APPLICANT: Burgess, Catherine E.,
; APPLICANT: Chant, John S.,
; APPLICANT: Chaudhuri, Amitabha,
; APPLICANT: DiPippo, Vincent A.,
; APPLICANT: Edinger, Shlomit R.,
; APPLICANT: Eisen, Andrew,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gangolli, Basha A.,
; APPLICANT: Gorman, Linda,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Khrantsov, Nikolai,
; APPLICANT: Li, Li,
; APPLICANT: Malyankar, Uriel M.,
; APPLICANT: MacDougall, John R.,
; APPLICANT: Mezes, Peter S.,
; APPLICANT: Miller, Charles E.,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Ooi, Chean Eng,
; APPLICANT: Ort, Tatiana,
; APPLICANT: Padigaru, Muralidhara,
; APPLICANT: Paturajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Rieger, Daniel K.,
; APPLICANT: Rothenberg, Mark E.,
; APPLICANT: Shenoy, Sureen G.,
; APPLICANT: Spaderna, Steven K.,
; APPLICANT: Spytek, Kimberley A.,
; APPLICANT: Taupier, Jr., Raymond J.,
; APPLICANT: Vernet, Corine A.M.,
; APPLICANT: Zernhusen, Bryan D.,
; APPLICANT: Zhong, Wei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27

```

```

; PRIOR APPLICATION NUMBER: 60/364,000
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,409
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/364,227
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/334,027
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/331,641
; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See File Wrapper or PLM.
; NUMBER OF SEQ ID NOS: 673
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 304
; LENGTH: 4455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-226-304

```

```

Query Match          76.6%; Score 36; DB 15; Length 4455;
Best Local Similarity 100.0%; Pred. No. 7.9e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GPGTK 6
        |||||
Db      2503 GPGTK 2508

```

```

Search completed: April 18, 2005, 09:04:05
Job time : 31.927 secs

```

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 07:32:05 ; Search time 8.67153 Seconds

(without alignments)  
99.861 Million cell updates/sec

Title: US-10-652-334-1

Perfect score: 47

Sequence: 1 GPPGKTKYT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	83.0	369	T29836	hypothetical prote
2	36	76.6	312	AH0631	hypothetical prote
3	36	76.6	1546	CGH02E	collagen alpha 2(I)
4	36	76.6	2405	T08164	dynam alpha heavy
5	36	76.6	4466	S17231	dynam beta heavy
6	36	76.6	4466	S17653	dynam beta heavy
7	35.5	75.5	319	C69507	activator 1, repli
8	35.5	75.5	319	B86350	hypothetical prote
9	35.5	75.5	340	A36988	replication factor
10	35.5	75.5	342	T43410	replication factor
11	35.5	75.5	353	S45531	replication factor
12	35.5	75.5	363	A45253	activator 1 37k ch
13	35.5	75.5	369	B84356	AAA-type ATPase [i
14	35.5	75.5	369	A96804	probable replicat
15	35.5	75.5	409	E69309	ATPase AAA homolog
16	35.5	75.5	447	S44809	DNA helicase - Aqu
17	35.5	75.5	530	D70476	DNA helicase - Aqu
18	35.5	75.5	619	E70597	hypothetical prote
19	35.5	75.5	635	T02699	probable helicase
20	35.5	75.5	989	T48845	inulin II gene en
21	35.5	75.5	993	A47500	Ig mu chain switch
22	35.5	75.5	993	S35633	DNA-binding protei
23	35.5	75.5	1121	S30862	DNA dependent Arpa
24	35	74.5	180	C69487	cyclidylate kinase
25	35	74.5	469	G84779	hypothetical prote
26	35	74.5	668	C75264	probable serine/th
27	35	74.5	1053	S46199	probable complemen
28	35	74.5	4006	T09070	probable tenascin
29	35	74.5	4135	T42629	tenascin-X - bovin

30	34.5	73.4	448	2	AF2198	AAA superfamily AT
31	34.5	73.4	555	2	B71420	hypothetical prote
32	34.5	73.4	613	2	AD2086	hypothetical prote
33	34.5	73.4	633	2	T28788	hypothetical prote
34	34.5	73.4	660	2	T41580	probable dna-bindi
35	34.5	73.4	663	2	H64312	probable DNA helic
36	34.5	73.4	678	2	T42668	hypothetical prote
37	34.5	73.4	683	2	S34700	probable purine nu
38	34.5	73.4	692	2	B90113	hypothetical prote
39	34.5	73.4	711	2	G85610	hypothetical prote
40	34.5	73.4	751	2	C84367	DNA binding protei
41	34.5	73.4	821	2	C84304	DNA helicase (limp
42	34.5	73.4	935	2	S62476	hypothetical prote
43	34.5	73.4	971	2	S23408	prematurely termin
44	34.5	73.4	1004	2	A39611	probable GTP-bind
45	34.5	73.4	1048	2	C86189	protein T25N20.11

#### ALIGNMENTS

RESULT 1  
T29836  
Hypothetical protein B0222.5 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T29836  
R:Du, Z.; Gattung, S.  
Submitted to the EMBL Data Library, February 1996  
A:Description: The sequence of C. elegans cosmid B0222.  
A:Reference number: Z20696  
A:Accession: T29836  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-369 <DUZ>  
A:Cross-references: UNIPROT:Q17457; EMBL:U50312; PDB:AAA92322.1; CSDP:B0222.5  
A:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CSDP:B0222.5  
A:Introns: 27/2; 47/1; 86/1; 114/1; 158/1; 206/3; 244/1; 275/1; 313/1

Query Match 83.0%; Score 39; DB 2; Length 369;  
Best Local Similarity 77.8%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 GPPGKTKYT 9  
Db 272 GPPGKTKYT 280

RESULT 2  
AH0631  
Hypothetical protein STY1143 [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhimurium  
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhimurium  
A:Name: this species has also been called *Salmonella typhi*  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C:Accession: AH0631  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Garra, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhimurium  
A:Reference number: AB0502; MIMD:21534947; PMID:11677608  
A:Accession: AH0631  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-312 <PAR>  
A:Cross-references: GB:AL513382; PDB:CAD08235.1; PDB:916502282; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY1143

Query Match 76.6%; Score 36; DB 2; Length 312;

Best Local Similarity 66.7%; Pred. No. 53;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPGTKXT 9  
Db 122 GPGTKAMT 130

## RESULT 3

CGH02E  
collagen alpha 2(XI) chain precursor - human (fragment)  
N:Alternate names: procollagen alpha 2(XI) chain  
N:Contains: proline/arginine-rich protein (PARP)  
C:Species: Homo sapiens (man)  
C:Date: 07-Jun-1990 #sequence\_revision 03-Oct-1995 #text\_change 22-Jun-1999  
C:Accession: S34790; A32645  
R:Zhukova, N.I.; Brewton, R.G.; Mayne, R.  
FEBS Lett. 326, 25-28, 1993  
A:Title: Molecular cloning of PARP (proline/arginine-rich protein) from human cartilage  
A:Reference number: S34790; MUID:93314796; PMID:8325374  
A:Accession: S34790  
A:Molecule type: mRNA  
A:Residues: 1-663 <ZHI>

A:Cross-references: EMBL:L18987; NID:G306439; PIDN:AA35498.1; PID:G306440  
R:Kimura, T.; Cheah, K.S.E.; Chan, S.D.H.; Lui, V.C.H.; Mattei, M.G.; van der Rest, M.;  
J. Biol. Chem. 264, 13910-13916, 1989  
A:Title: The human alpha2(XI) collagen (COL1A2) chain. Molecular cloning of cDNA and ge  
A:Reference number: A32645; MUID:89340485; PMID:2760050  
A:Accession: A32645  
A:Molecule type: DNA; mRNA  
A:Residues: 586-1546 <KIM>

A:Cross-references: GB:J04974; NID:G180714; PIDN:AA52034.1; PID:G180715  
A:Note: parts of this sequence were determined by protein sequencing  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
ed and subsequently O-glycosylated.  
C:Genetics:

A:Gene: GDB:COL1A2  
A:Cross-references: GDB:119788; OMIM:120290  
A:Map position: 6p21.3-6p21.3

A:Introns: 1302/3; 1320/3; 1332/3; 1350/3; 1440/1; 1477/3  
A:Note: the list of introns is incomplete

C:Complex: type XI collagen may be a heterotrimer of two alpha 1(XI) chains (see PIR:CGH  
3(XI) chain (see PIR:CGH6C). Initially linked by disulfide bonds among their carboxyl-  
med with desmosine cross-links made from lysine and allysine residues  
C:Function:

A:Description: structural component of extracellular fibrous polymer associated with cell  
A:Note: may play a role in controlling the lateral growth of collagen II fibrils  
C:Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology  
C:Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproli  
F:1-354/Domin: non-collagenous (fragment) #status predicted <NC3>  
F:1-187/Product: proline/arginine-rich PARP protein (fragment) #status predicted <PAR>  
F:255-305/Domin: collagenous, triple helix #status predicted <COL2>  
F:306-342/Domin: non-collagenous #status predicted <NC2>

F:429-431/Region: cell attachment (R-G-D) motif  
F:447-449/Region: cell attachment (R-G-D) motif  
F:1357-1359/Region: cell attachment (R-G-D) motif  
F:1357-1380/Region: carboxyl-terminal nonhelical telopeptide  
F:1381-1546/Domin: carboxyl-terminal propeptide (fragment) #status predicted <CTP>  
F:1403-1546/Domin: fibrillar collagen carboxyl-terminal homology (fragment) #status acty  
F:109-163, 1511-1546/Disulfide bonds: #status predicted  
F:319/Modified site: allysine (Lys) #status predicted  
F:426/1266/Modified site: 5-hydroxylysine (Lys) #status predicted  
F:426/1266/Binding site: carboxylate (Lys) (covalent) #status predicted  
F:927, 933, 1008, 1011, 1035, 1038, 1290, 1296, 1305, 1317, 1320/Modified site: 4-hydroxyproline  
F:929/Modified site: 4-hydroxyproline (Pro) #status atypical  
F:942, 1023, 1299/Modified site: 5-hydroxylysine (Lys) #status experimental  
F:942, 1023, 1299/Binding site: carboxylate (Lys) (covalent) #status experimental  
F:1427, 1433, 1450, 1459/Disulfide bonds: interchain #status predicted  
F:1460/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 76.6%; Score 36; DB 1; Length 1546;

Best Local Similarity 100.0%; Pred. No. 2,7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGTK 6  
Db 553 GPGTK 558

## RESULT 4

T08164  
dynein alpha heavy chain - Chlamydomonas reinhardtii (fragment)  
C:Species: Chlamydomonas reinhardtii  
C:Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 02-Feb-2001  
C:Accession: T08164  
R:Mitchell, D.R.; Brown, K.S.  
J. Cell Sci. 107, 635-644, 1994  
A:Title: Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes.  
A:Reference number: 216302; MUID:94274778; PMID:8006077  
A:Accession: T08164  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2405 <MIT>  
A:Cross-references: EMBL:L26049; NID:G415679; PIDN:AA57316.1; PID:G603079  
A:Experimental source: strain 21gr  
C:Genetics:

A:Note: Intron positions not resolved (incomplete sequence)  
C:Superfamily: dynein heavy chain, ciliary  
C:Keywords: nucleotide binding; P-loop  
F:575-582/Region: nucleotide-binding motif A (P-loop)

Query Match 76.6%; Score 36; DB 2; Length 2405;  
Best Local Similarity 100.0%; Pred. No. 4,2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGTK 6  
Db 1283 GPGTK 1288

## RESULT 5

S17231  
dynein beta heavy chain, ciliary - sea urchin (anthodidaris crassispina)  
N:Contains: dynein ATPase (EC 3.6.4.2)  
C:Species: Anthodidaris crassispina  
C:Date: 30-Sep-1991 #sequence\_revision 02-May-1994 #text\_change 09-Jul-2004  
C:Accession: S17231; PS0415  
R:Ogawa, K.  
Nature 352, 643-645, 1991

A:Title: Four ATP-binding sites in the midregion of the beta heavy chain of dynein.  
A:Reference number: S17231; MUID:91326104; PMID:1830928  
A:Accession: S17231  
A:Molecule type: mRNA  
A:Residues: 1-4466 <OGA>

A:Cross-references: UNIPROT:P39057; GB:D01021; NID:G2127202; PIDN:BA00827.1; PID:G2127203  
R:Ogawa, K.  
Proc. Jpn. Acad. B Phys. Biol. Sci. 67, 27-31, 1991

A:Title: ATP-binding site in dynein beta-heavy chain: identification by molecular cloning  
A:Reference number: PS0415  
A:Accession: PS0415  
A:Molecule type: mRNA  
A:Residues: 764-1001, 'APC', 1005-2036, 'VPSSVER' <OG2>

C:Superfamily: dynein heavy chain, ciliary  
C:Keywords: ATP, heterotrimer; hydrolyase; microtubule binding; nucleotide binding; P-loop  
F:154-161/Region: nucleotide-binding motif A (P-loop)  
F:1854-1859/Region: nucleotide-binding motif A (P-loop)  
F:2133-2140/Region: nucleotide-binding motif A (P-loop)  
F:2460-2467/Region: nucleotide-binding motif A (P-loop)  
F:2805-2812/Region: nucleotide-binding motif A (P-loop)  
F:160/Binding site: ATP (Lys) #status predicted  
F:1858/Binding site: ATP (Lys) #status predicted  
F:2133/Binding site: ATP (Lys) #status predicted  
F:2466/Binding site: ATP (Lys) #status predicted

F:2811/Binding site: ATP (Lys) #status predicted

Query Match 76.6%; Score 36; DB 1; Length 446;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGTK 6  
Db 2514 GPGTK 2519

#### RESULT 6

S17653

dynamin beta heavy chain, ciliary - sea urchin (*Tridpneustes gratilla*)

C:Species: *Tridpneustes gratilla*

C:Date: 04-Dec-1992 #sequence\_revision 02-May-1994 #text\_change 09-Jul-2004

C:Accession: S17653; S24628

R:Gibbons, I.R.; Gibbons, B.H.; Mocq, G.; Asal, D.J.

Nature 352, 640-643, 1991

A:Title: Multiple nucleotide-binding sites in the sequence of dynamin beta heavy chain.

A:Reference number: S17653; MUID:91326103; PMID:1830927

A:Accession: S17653

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Accession: S17653

A:Cross-references: UNIPROT:P23098; EMBL:X59603; NID:g10709; PIDN:CAA42170.1; PID:g10710

A:Molecule type: protein

A:Accession: S24628

A:Reference: 162-172;1193-1204;3240-3259;3325-3339 <GIB2>

C:Superfamily: dynein heavy chain, ciliary

C:Keywords: ATP; heterotrimer; hydrolyase; microtubule binding; nucleotide binding; P-1

P:154-161/Region: nucleotide-binding motif A (P-loop)

P:1852-1859/Region: nucleotide-binding motif A (P-loop)

P:2133-2140/Region: nucleotide-binding motif A (P-loop)

P:2460-2467/Region: nucleotide-binding motif A (P-loop)

P:2805-2812/Region: nucleotide-binding motif A (P-loop)

P:1858/Binding site: ATP (Lys) #status predicted

P:2139/Binding site: ATP (Lys) #status predicted

P:2466/Binding site: ATP (Lys) #status predicted

F:2811/Binding site: ATP (Lys) #status predicted

Query Match 76.6%; Score 36; DB 1; Length 446;

Best Local Similarity 100.0%; Pred. No. 7.8e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGTK 6

Db 2514 GPGTK 2519

#### RESULT 7

C69507

activator 1, replication factor C, 35 KD subunit homolog - *Archaeoglobus fulgidus*

C:Species: *Archaeoglobus fulgidus*

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C:Accession: C69507

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.B.; Ketchum, K.A.; Dodson, R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirtress, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Ueberbeck, T.; Cotton, M.D.; Spriggs, T.; Attlich, P.; Kaine, B.F.; Sykes, S. Smith, H.O.; Moese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea

A:Reference number: A69250; MUID:9804343; PMID:9389475

A:Accession: C69507

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Accession: S17653 <KIR>

A:Cross-references: UNIPROT:O28219; GB:AE000961; GB:AE000782; NID:g2689284; PIDN:AA88918

C:Superfamily: phage T4 DNA polymerase accessory protein 44

Query Match 75.5%; Score 35.5; DB 2; Length 319;

Best Local Similarity 80.0%; Pred. No. 66;  
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 GPGTK-TXT 9  
Db 45 GPGTK-TXT 54

#### RESULT 8

B86350

hypothetical protein P8K7.11 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: B86350

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federapfel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maitl, R.; Marziani, R.

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, J.

A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: B86350

A:Status: preliminary

A:Molecule type: DNA

A:Accession: S17653 <STO>

A:Cross-references: UNIPROT:Q9X109; GB:AE005172; NID:g5263320; PIDN:AA41422.1; GSPDB:GN

C:Genetics:

A:Map position: 1

C:Superfamily: phage T4 DNA polymerase accessory protein 44

Query Match 75.5%; Score 35.5; DB 2; Length 319;

Best Local Similarity 80.0%; Pred. No. 66;

Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 GPGTK-TXT 9

Db 49 GPGTK-TXT 58

#### RESULT 9

A36988

replication factor C chain RFC3 [validated] - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: protein N0533; protein YNL0533; protein YNL290w

C:Species: *Saccharomyces cerevisiae*

C:Date: 25-Aug-1995 #sequence\_revision 25-Aug-1995 #text\_change 09-Jul-2004

C:Accession: A36988; S59666; S60412; S63264; S63266; S44762

R:Li, X.; Burgers, P.M.J.

Proc. Natl. Acad. Sci. U.S.A. 91, 868-872, 1994

A:Title: Molecular cloning and expression of the *Saccharomyces cerevisiae* RFC3 gene, an

A:Reference number: A36988; MUID:94134732; PMID:8302859

A:Accession: A36988

A:Status: preliminary

A:Molecule type: DNA

A:Accession: S17653 <LTA>

A:Cross-references: UNIPROT:P36629; EMBL:L16755; NID:g4339117; PIDN:AAA34969.1; PID:g43391

R:Callimann, G.; Flen, K.; Kobayashi, R.; Stillman, B.

Mol. Cell. Biol. 15, 4661-4671, 1995

A:Title: Characterization of the five replication factor C genes of *Saccharomyces cerevi*

A:Reference number: S59664; MUID:95379800; PMID:7651383

A:Accession: S59666

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Accession: S17653 <STO>

A:Cross-references: EMBL:U26029; NID:g841465; PIDN:AA449062.1; PID:g841466

R:Maurel, K.C.T.; Urbansky, J.H.M.; Planca, R.J.

Yeast 11, 1303-1310, 1995

A:Title: Sequence analysis of a 30 kb DNA segment from yeast chromosome XIV carrying a r

C, and a novel putative serine/threonine protein kinase gene.

A:Reference number: S60394; MUID:96132033; PMID:8553702

A:Accession: S60412  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Reference number: 221749  
 A:Accession: T38458  
 A:Molecule type: DNA  
 A:Residues: 1-340 <MAN>  
 A:Cross-references: EMBL:U23084; NID:G1050853; PIDD:AA049110.1; PID:G1050872  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1995  
 R:Mesenguy, F.; Dubois, B.; Vierendeels, F.; Scherens, B.; Pierard, A.; Glansdorff, N.  
 submitted to the Protein Sequence Database, April 1996  
 A:Reference number: S63245  
 A:Accession: S63264  
 A:Molecule type: DNA  
 A:Residues: 1-340 <MES>  
 A:Cross-references: EMBL:Z71566; NID:G1302374; PIDD:CAA96207.1; PID:G1302375; GSPDB:GN00  
 R:Maurer, C.T.C.; Urbanus, J.H.M.; Planta, R.J.  
 submitted to the Protein Sequence Database, April 1996  
 A:Reference number: S63266  
 A:Accession: S63266  
 A:Molecule type: DNA  
 A:Residues: 1-340 <MAN>  
 A:Cross-references: EMBL:Z71566; NID:G1302374; PIDD:CAA96207.1; PID:G1302375; GSPDB:GN00  
 A:Experimental source: strain S288C  
 C:Genetics:  
 A:Gene: SGD:RPC3; MIPS:YML290W  
 A:Cross-references: SGD:S0005234; MIPS:YML290W  
 A:Map position: 14L  
 C:Complex: heteropentamer consists of RFC1 (PIR:S44763), RFC2 (PIR:S45531), RFC3 (PIR:A3  
 C:Function:  
 A:Description: the RFC complex functions as a structure-specific, DNA-dependent ATPase,  
 A>Note: each of the five chains are essential for cell proliferation  
 C:Superfamily: phase T4 DNA polymerase accessory protein 44  
 C:Keywords: DNA binding

Query Match 75.5%; Score 35.5; DB 2; Length 340;  
 Best Local Similarity 80.0%; Pred. No. 70;  
 Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Db 53 GPGGTGKTST 62

RESULT 10

T43410  
 replication factor C chain RFC3 - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
 C:Accession: T43410; T43412; T43651; T38458  
 R:Nojima, H.  
 submitted to the EMBL Data Library, August 1998  
 A:Description: Rfc3 of S. pombe.  
 A:Reference number: Z22494  
 A:Accession: T43410  
 A:Molecule type: DNA  
 A:Residues: 1-342 <NOS>  
 A:Cross-references: UNIPROT:O14003; EMBL:AB017039; PIDD:BA082745.1  
 R:Nojima, H.  
 submitted to the EMBL Data Library, August 1998  
 A:Description: CDNA of rfc3.  
 A:Reference number: Z22496  
 A:Accession: T43412  
 A:Molecule type: mRNA  
 A:Residues: 1-342 <NOS>  
 A:Cross-references: EMBL:AB017040; PIDD:BA082746.1  
 R:Gray, F.C.; MacNeill, S.A.  
 submitted to the EMBL Data Library, November 1998  
 A:Description: Characterisation of the rfc3+ gene encoding a subunit of replication fact  
 A:Reference number: Z22600  
 A:Accession: T43651  
 A:Molecule type: DNA  
 A:Residues: 1-342 <GRA>  
 A:Cross-references: EMBL:AJ012839; PIDD:CA038106.1  
 R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, September 1997  
 A:Reference number: 221749  
 A:Accession: T38458  
 A:Molecule type: DNA  
 A:Residues: 119-342 <MDU>  
 A:Cross-references: EMBL:Z98978; PIDD:CA039134.1; GSPDB:GN000066; SPDB:SPAC27E2.10C  
 A:Experimental source: strain 972n-; cosmid C27E2  
 C:Genetics:  
 A:Gene: rfc3; SPDB:SPAC27E2.10C  
 A:Map position: 1  
 A:Insertions: 25/2; 47/1; 118/2; 143/1; 267/1  
 C:Superfamily: phase T4 DNA polymerase accessory protein 44

Query Match 75.5%; Score 35.5; DB 2; Length 342;  
 Best Local Similarity 80.0%; Pred. No. 71;  
 Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Db 63 GPGGTGKTST 72

RESULT 11

S45531  
 replication factor C chain RFC2 - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein J1808; protein YJR068W  
 C:Species: Saccharomyces cerevisiae  
 C>Date: 31-Mar-1992 #sequence\_revision 14-Sep-1994 #text\_change 09-Jul-2004  
 C:Accession: S45531; S57087; S59665; S59879; S71690  
 R:Noskov, V.; Mak, S.; Kawasaki, Y.; Leem, S.H.; Ono, B.I.; Araki, H.; Pavlov, Y.; Sugit  
 Nucleic Acids Res. 22, 1527-1535, 1994  
 A>Title: The RFC2 gene encoding a subunit of replication factor C of Saccharomyces cerevi  
 A:Reference number: S45531; MUID:94261414; PMID:8202350  
 A:Accession: S45531  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-353 <NOS>  
 A:Cross-references: UNIPROT:P40348; EMBL:D28499; NID:G498462; PIDD:BA05858.1; PID:G49846  
 R:Mann, V.; Huang, M.B.; Galibert, F.  
 submitted to the Protein Sequence Database, September 1995  
 A:Reference number: S57085  
 A:Accession: S57087  
 A:Molecule type: DNA  
 A:Residues: 1-353 <MAN>  
 A:Cross-references: EMBL:Z49568; NID:G1015746; PIDD:CA089596.1; PID:G1015747; MIPS:YJR068  
 R:Cullmann, G.; Fien, K.; Kobayashi, R.; Stillman, B.  
 Mol. Cell. Biol. 15, 4661-4671, 1995  
 A>Title: Characterization of the five replication factor C genes of Saccharomyces cerevisi  
 A:Reference number: S59664; MUID:95379808; PMID:7651383  
 A:Accession: S59665  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-353 <CUH>  
 A:Cross-references: EMBL:U26028; NID:G841463; PIDD:AA049061.1; PID:G841464  
 A:Accession: S59979  
 A:Molecule type: protein  
 A:Residues: 55-71;187-190;296-315 <CUW>  
 R:Huang, M.B.; Mann, V.; Chuat, J.C.; Galibert, F.  
 yeast 12, 869-875, 1996  
 A>Title: Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open reading frames  
 A:Reference number: S71676; MUID:96437976; PMID:8840504  
 A:Accession: S71690  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-353 <HUA>  
 A:Cross-references: EMBL:L47993; NID:G1019675; PIDD:AB039294.1; PID:G1019690  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995  
 C:Genetics:  
 A:Gene: SGD:RPC2  
 A:Cross-references: SGD:S0003829; MIPS:YJR068W  
 A:Map position: 10R  
 C:Superfamily: phase T4 DNA polymerase accessory protein 44  
 C:Keywords: DNA binding; nucleus



```

Query Match      75.5%; Score 35.5; DB 2; Length 353;
Best Local Similarity 80.0%; Pred. No. 73;
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GPGGT-KTXXT 9
DB 65 GPGGTGKTST 74

RESULT 12
Accession: A45253
activator 1 37K chain - human
N/Alternate names: replication factor C 37K chain
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Aug-1998
C/Accession: A45253; B45253
R/Chen, M.; Pan, Z.Q.; Hurwitz, J.
Proc. Natl. Acad. Sci. U.S.A. 89, 5211-5215, 1992
A/Title: Studies of the cloned 37-kDa subunit of activator 1 (replication factor C) of H
A/Reference number: A45253; MUID:9230215; PMID:1351677
A/Accession: A45253
A/Molecule type: mRNA
A/Residues: 1-363 <CHR>
A/Experimental source: HeLa cells
A/Note: sequence extracted from NCBI backbone (NCBIN:106894, NCBI:P:106895)
A/Accession: B45253
A/Molecule type: protein
A/Residues: 65-83;94-101;125-134;147-163;261-272 <CH2>
C/Superfamily: phage T4 DNA polymerase accessory protein 44

Query Match      75.5%; Score 35.5; DB 2; Length 363;
Best Local Similarity 80.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GPGGT-KTXXT 9
DB 78 GPGGTGKTST 87

RESULT 13
Accession: B84356
AAA-type ATPase [imported] - Halobacterium sp. NRC-1
C/Species: Halobacterium sp. NRC-1
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: B84356
R/Ng, W.V.; Kennedy, S.P.; Mahatras, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leitbauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
A/Title: Genome sequence of Halobacterium species NRC-1.
A/Reference number: A84160; MUID:20504483; PMID:11016950
A/Accession: B84356
A/Molecule type: DNA
A/Status: preliminary
A/Residues: 1-369 <STO>
A/Cross-references: UNIPROT:Q9HNK6; GB:AE004437; NID:910581481; PIDN:AAG20214.1; GSPDB:G
A/Genetics:
C/Superfamily: ATP-dependent 26S protease; FtsH/SEC18/CDC48-type ATP-binding domain ho

Query Match      75.5%; Score 35.5; DB 2; Length 369;
Best Local Similarity 80.0%; Pred. No. 76;
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GPGGT-KTXXT 9
DB 140 GPGGTGKTST 149

RESULT 14

```

```

probable replication factor C, 24844-22715 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: A96804
R/Ritholgis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzfall,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: A96804
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-369 <STO>
A/Cross-references: UNIPROT:Q9CA08; GB:AE005173; NID:96382495; PIDN:AA07781.1; GSPDB:GN
C/Genetics:
A/Map position: 1
A/Superfamily: phage T4 DNA polymerase accessory protein 44

Query Match      75.5%; Score 35.5; DB 2; Length 369;
Best Local Similarity 80.0%; Pred. No. 76;
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GPGGT-KTXXT 9
DB 79 GPGGTGKTST 88

RESULT 15
Accession: B69309
ATPase AAA homolog - Archaeoglobus fulgidus
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C/Accession: B69309
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glock, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A/Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Moese, C.R.; Venter, J.C.
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A/Reference number: A69250; MUID:98049343; PMID:9389475
A/Accession: B69309
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-409 <KLR>
A/Cross-references: UNIPROT:Q29773; GB:AE001071; GB:AE000782; NID:92689394; PIDN:AAB9076
C/Superfamily: ATP-dependent 26S protease; FtsH/SEC18/CDC48-type ATP-binding domain ho
C/Keywords: ATP; nucleotide binding; P-loop
F:178-382/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAMP>
F:205-212/Region: nucleotide-binding motif A (P-loop)

Query Match      75.5%; Score 35.5; DB 2; Length 409;
Best Local Similarity 80.0%; Pred. No. 85;
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GPGGT-KTXXT 9
DB 205 GPGGTGKTST 214

Search completed: April 18, 2005, 08:05:57
Job time : 10.6715 secs

```

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using BW model

Run on: April 18, 2005, 07:33:55 ; Search time 37.5766 Seconds  
(without alignments)  
122.648 Million cell updates/sec

Title: US-10-652-334-1  
Perfect score: 47  
Sequence: 1 GPGTKXTX 9

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 03:.\*  
1: uniprot\_sprot:.\*  
2: uniprot\_trembl:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	39	83.0	Q17457	Q17457 caenorhabd1
2	38	80.9	Q9LIA0	Q9LIA0 streptomyc
3	37	78.7	Q7UAK1	Q7UAK1 shigella fl
4	37	78.7	Q831H9	Q831H9 shigella fl
5	37	78.7	Q831H9	Q831H9 shigella fl
6	37	78.7	Q831H9	Q831H9 shigella fl
7	37	78.7	Q831H9	Q831H9 shigella fl
8	37	78.7	Q831H9	Q831H9 shigella fl
9	37	78.7	Q831H9	Q831H9 shigella fl
10	37	78.7	Q831H9	Q831H9 shigella fl
11	37	78.7	Q831H9	Q831H9 shigella fl
12	37	78.7	Q831H9	Q831H9 shigella fl
13	37	78.7	Q831H9	Q831H9 shigella fl
14	37	78.7	Q831H9	Q831H9 shigella fl
15	37	78.7	Q831H9	Q831H9 shigella fl
16	37	78.7	Q831H9	Q831H9 shigella fl
17	37	78.7	Q831H9	Q831H9 shigella fl
18	37	78.7	Q831H9	Q831H9 shigella fl
19	37	78.7	Q831H9	Q831H9 shigella fl
20	37	78.7	Q831H9	Q831H9 shigella fl
21	37	78.7	Q831H9	Q831H9 shigella fl
22	37	78.7	Q831H9	Q831H9 shigella fl
23	37	78.7	Q831H9	Q831H9 shigella fl
24	37	78.7	Q831H9	Q831H9 shigella fl
25	37	78.7	Q831H9	Q831H9 shigella fl
26	37	78.7	Q831H9	Q831H9 shigella fl
27	37	78.7	Q831H9	Q831H9 shigella fl
28	37	78.7	Q831H9	Q831H9 shigella fl
29	37	78.7	Q831H9	Q831H9 shigella fl
30	37	78.7	Q831H9	Q831H9 shigella fl
31	37	78.7	Q831H9	Q831H9 shigella fl

32	36	76.6	1736	1	CA2B_MOUSE	Q64739 mus musculu
33	36	76.6	1827	2	Q8UUM5	Q8UUM5 oryzias lat
34	36	76.6	1860	2	Q81ZC6	Q81ZC6 homo sapien
35	36	76.6	1863	2	Q731L5	Q731L5 homo sapien
36	36	76.6	4466	1	DYHC_ANTCR	P39057 antioctidari
37	36	76.6	4466	1	DYHC_TRIGR	P23098 tritpneustes
38	36	76.6	4488	2	DYHA_CHLRE	Q39610 chlamydomon
39	36	76.6	4499	2	DYHA_CHLRE	Q39610 chlamydomon
40	36	76.6	4511	2	Q7PXZ5	Q7PXZ5 anophelis g
41	35.5	75.5	202	2	Q7TC93	Q7TC93 brachydanti
42	35.5	75.5	319	1	RFC5_ARCFU	Q28219 archaeglob
43	35.5	75.5	319	2	Q9X109	Q9X109 archidopais
44	35.5	75.5	326	2	Q8QNA5	Q8QNA5 ectocarpus
45	35.5	75.5	329	2	Q6CNE4	Q6CNE4 kluyveromyc

## ALIGNMENTS

RESULT 1  
Q17457 PRELIMINARY; PRT: 369 AA.  
ID Q17457  
AC Q17457  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Hypothetical protein B0222.5.  
GN Name=B0222.5; ORFNames=B0222.5;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RG WormBase Consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Du Z., Gattung S.;  
RT "The sequence of C. elegans cosmid B0222.";  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Wilson R.;  
RT Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RG WormBase Consortium;  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
CC -1- SMILIRITY: Contains 1 BPTI/Kunitz inhibitor domain.  
DR EMBL: U50312; AAA92322.1; -.  
DR PIR: T29836; T29836.  
DR HSSP: P31713; ISHP.  
DR WormBase: WBGene00105056; B0222.5.  
DR WormPeP: B0222.5; CS06696.  
DR GO: GO:0004867; F:serine-type endopeptidase inhibitor activity; IEA.  
DR InterPro: IPR002223; Prot\_Inh\_Kunz-m.  
DR InterPro: IPR000716; Thyroglobulin\_1.  
DR Pfam: PF00014; Kunitz\_BPTI\_1.  
DR Pfam: PF00086; Thyroglobulin\_1; 1.  
DR ProDom: PD000222; Prot\_Inh\_Kunz-m; 1.

DR SMART; SM00131; KU; 1.  
 DR SMART; SM00211; TY; 2.  
 DR PROSITE; PS00280; BPTI\_KUNITZ\_1; UNKNOWN\_1.  
 DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
 DR PROSITE; PS00484; TRYROGLOBULIN\_1; UNKNOWN\_1.  
 DR Hypothetical protein.  
 SQ SEQUENCE 369 AA; 41135 MW; BEDDE119B96969BC CRC64;

Query Match 83.0%; Score 39; DB 2; Length 369;  
 Best Local Similarity 77.8%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPGTKTXXT 9  
 |||||  
 Db 272 GPGTKTXXT 280

## RESULT 2

Q9L1A0 PRELIMINARY; PRT; 264 AA.  
 ID 09L1A0  
 AC 09L1A0  
 DT 01-OCT-2000 (TRENBLrel. 15, Created)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE Putative acetyltransferase.  
 GN ORFNames=SC10G8..02;  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteriae; Actinobacteriales;  
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.  
 CX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleiser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleiser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,  
 RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J.R., Barrall B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2).";  
 RL Nature 417:141-147(2002).

DR EMBL; AL393131; CAB76276.1; -.  
 DR GO; GO:0008080; F:N-acetyltransferase activity; IEA.  
 DR GO; GO:0016740; P:transferase activity; IEA.  
 DR InterPro; IPR000182; GCSAcetyl trams.  
 DR Pfam; PF00583; Acetyltransferase\_1; I.  
 KM Complete proteome; Transferase.  
 SQ SEQUENCE 264 AA; 28668 MW; 07259EBB96FC20B6 CRC64;

Query Match 80.9%; Score 38; DB 2; Length 264;  
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGTKTXXT 7  
 |||||  
 Db 113 GPGTKTXXT 119

## RESULT 3

Q7UAK1 PRELIMINARY; PRT; 255 AA.  
 ID 07UAK1  
 AC 07UAK1  
 DT 01-OCT-2003 (TRENBLrel. 25, Created)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE Hypothetical protein.  
 OS OrderedLocustNames=S4646;  
 GN Shigella flexneri.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Shigella.  
 CX NCBI\_TaxID=623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=2457T;  
 RX MEDLINE=22590274; PubMed=12704152;  
 RX DOI=10.1128/IAI.71.5.2775-2786.2003;  
 RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,  
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,  
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,  
 RA Schwartz D.C., Blattner F.R.;  
 RT "Complete genome sequence and comparative genomics of Shigella  
 flexneri serotype 2a strain 2457T.";  
 RL Infect. Immun. 71:2775-2786(2003).  
 DR EMBL; AE016993; AAP19570.1; -.  
 DR InterPro; IPR001279; Bactamase-like.  
 DR Pfam; PF00753; Lactamase\_B; 1.  
 KM Hypothetical protein.  
 SQ SEQUENCE 255 AA; 27898 MW; 9CBA06FBC1BE6E2B3 CRC64;

Query Match 78.7%; Score 37; DB 2; Length 255;  
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPGTKTXXT 9  
 |||||  
 Db 71 GPGTKTXXT 79

## RESULT 4

O83IH9 PRELIMINARY; PRT; 276 AA.  
 ID 083IH9  
 AC 083IH9  
 DT 01-JUN-2003 (TRENBLrel. 24, Created)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Hypothetical protein SF4376.  
 GN OrderedLocustNames=SF4376;  
 OS Shigella flexneri.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Shigella.  
 CX NCBI\_TaxID=623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=301 / Serotype 2a;  
 RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;  
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
 RA Yang Q., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
 RA Sun L., Xue Y., Zhao A., Gao Y., Kan B., Ding K., Chen S.,  
 RA Chang H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
 RA Yu J.;  
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity  
 through comparison with genomes of Escherichia coli K12 and O157.";  
 RL Nucleic Acids Res. 30:4432-4441(2002).  
 DR EMBL; AE015443; AAN45792.1; -.  
 DR InterPro; IPR001279; Bactamase-like.  
 DR Pfam; PF00753; Lactamase\_B; 1.  
 KM Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 276 AA; 30087 MW; 8C1F192E049E1145 CRC64;

Query Match 78.7%; Score 37; DB 2; Length 276;  
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPGTKTXXT 9  
 |||||  
 Db 92 GPGTKTXXT 100

## RESULT 5

O83IE3 PRELIMINARY; PRT; 323 AA.  
 ID 083IE3

```

AC 0831E3;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DE Putative Iron-siderophore binding lipoprotein.
GN OrderedLocustNames=TW066;
OS Tropheryma whippelii (strain TW08/27) (Whipple's bacillus).
OC Bacteria; Actinobacteriales; Actinobacteridae; Actinomycetales;
OC Micrococccineae; Cellulomonadaceae; Tropheryma.
CX NCBI_TaxID=218496;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22495039; PubMed=12606174; DOI=10.1016/S0140-6736(03)12597-4;
RA Bentley S.D., Mawald M., Murphy L.D., Pallen M.J., Yeats C.A.,
RA Dover L.G., Norbertczak H.T., Beers G.S., Quail M.A., Harris D.E.,
RA von Herbay A., Goble A., Rutter S., Squares R., Squares S.,
RA Barrett B.G., Parkhill J., Reisman D.A.;
RT "Sequencing and analysis of the genome of the Whipple's disease
RT bacterium Tropheryma whippelii."
RL Lancet 361:637-644(2003).
DR EMBL, BX251410; CAD66753.1; -
DR GO; GO:0005381; F:Iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
KW Complete proteome.
SQ SEQUENCE 323 AA; 35377 MW; DPA3BD97A3F11FCD CRC64;

Query Match 78.7%; Score 37; DB 2; Length 323;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GPGTKT 7
Db 146 GPGTKS 152

RESULT 6
O83N44 PRELIMINARY; PRT; 323 AA.
ID O83N44;
AC O83N44;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DE Iron (III) diclitate ABC transporter substrate-binding protein.
GN Name=fcsh; OrderedLocustNames=TW056;
OS Tropheryma whippelii (strain Twist) (Whipple's bacillus).
OC Bacteria; Actinobacteriales; Actinobacteridae; Actinomycetales;
OC Micrococccineae; Cellulomonadaceae; Tropheryma.
CX NCBI_TaxID=203267;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Twist;
RX MEDLINE=22784088; PubMed=12902375;
RA Raoult D., Ogata H., Audic S., Robert C., Sühre K., Drancourt M.,
RA Claverie J.-M.;
RT "Tropheryma whippelii Twist: a human pathogenic Actinobacteria with a
RT reduced genome."
RL Genome Res. 13:1800-1809(2003).
DR EMBL, AE016850; AA044153.1; -
DR GO; GO:0005381; F:Iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
KW Complete proteome.
SQ SEQUENCE 323 AA; 35377 MW; DPA3BD97A3F11FCD CRC64;

Query Match 78.7%; Score 37; DB 2; Length 323;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GPGTKT 7
Db 146 GPGTKS 152

RESULT 7

```

```

O88AM5 PRELIMINARY; PRT; 438 AA.
ID O88AM5;
AC O88AM5;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DE Putative ABC transporter, periplasmic sugar-binding protein,
DE Sugar ABC transporter, periplasmic sugar-binding protein,
DE putative.
GN OrderedLocustNames=PSPT0364;
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
CX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA Buell C.R., Joardar V., Lindberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidson T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H.M., Pedrovica N.B., Tran B., Russell D., Berry K.J.,
RA Ueterbach T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bender C.L., White O., Fraser C.M., Collier A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
DR EMBL, AB016857; AA053908.1; -
DR TIGR; PSPT0364; -
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR006059; SBP_bac_1.
DR Pfam; PF01547; SBP_bac_1; 1.
KW Complete proteome.
SQ SEQUENCE 438 AA; 47925 MW; 9EA429C47F88B66 CRC64;

Query Match 78.7%; Score 37; DB 2; Length 438;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 PPGTKXT 9
Db 343 PPGTKST 350

RESULT 8
O6H329 PRELIMINARY; PRT; 532 AA.
ID O6H329;
AC O6H329;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Putative endo-beta-1,4-glucanase.
GN Name=OSJNOA018M17.6; Synonyms=B1339H09.3;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
CX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX Sasaki T., Matsumoto T., Katayose Y.;
RX Submitted (JUN-2004) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX Sasaki T., Matsumoto T., Katayose Y.;
RX Submitted (MAY-2003) to the EMBL/Genbank/DBJ databases.
RL EMBL, AP007149; BAD26550.1; -
DR EMBL, AP006453; BAD26493.1; -
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.

```

```

DR InterPro; IPR001701; Glyco_hydro_9.
DR InterPro; IPR008928; Glyco_trans_6hp.
DR Pfam; PF00759; Glyco_hydro_9; 1.
DR PROSITE; PS00592; GLYCOSYL_HYDROL_F9_1; 1.
SQ SEQUENCE 532 AA; 58570 MW; 594D747C1A4EDDE8 CRC64;

Query March
Best Local Similarity 78.7%; Score 37; DB 2; Length 532;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy 2 PEGTKTXXT 9
Db 12 PEGTKTXXT 19

RESULT 9
ID Q8BK01 PRELIMINARY; PRT; 577 AA.
AC Q8BK01;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein XAC2118.
GN OrderedLocustNames=XAC2118;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farsh C.S., Furman L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavaro F., Cardoso J., Chambergo F., Clapina L.P.,
RA Cleaveland R.M.B., Coutinho L.L., Curcio-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gulber A.,
RA Katsuyama A.M., Kiehl L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanki J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinoja L.A.F., Takita M.A., Tamura R.B., Teixeira B.C., Tezsa R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White P.F.,
RA Secubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
RW EMBL; AB011847; AAM36971.1; -
KW Complete proteome.
SQ SEQUENCE 577 AA; 60709 MW; 85B3E689C3E3632 CRC64;

Query Match
Best Local Similarity 78.7%; Score 37; DB 2; Length 577;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Cy 1 GPGTKTXXT 9
Db 460 GPGTKTXXT 468

RESULT 10
ID Q9LGA3 PRELIMINARY; PRT; 586 AA.
AC Q9LGA3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE ESTs D22655 (C0749).
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriophytaceae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC PUBMED=12447438; DOI=10.1038/nature01184;
RX Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nishimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Maekawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hishikita S., Honda M., Ichikawa Y., Ikonuma A., Iijima M., Ikeda M.,
RA Ikono M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tanji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshinaka R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1."
RL Nature 420:312-316(2002).
RW EMBL; AP002538; BAB03379.1; -
DR GRMeme; Q9LGA3; -
SQ SEQUENCE 586 AA; 63316 MW; 8431BED1092BDE0FB CRC64;

Query Match
Best Local Similarity 78.7%; Score 37; DB 2; Length 586;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 GPGTKTXXT 7
Db 413 GPGTKTXXT 419

RESULT 11
ID Q6A548;
AC Q6A548;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Telomerase reverse transcriptase (EC 2.7.7.-) (Telomerase catalytic
DE subunit).
GN Name=TERF;
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC PUBMED=15225880; DOI=10.1016/j.gene.2004.03.032;
RX Nasir L., Gault E., Campbell S., Veeramalai M., Gilbert D.,
RA McFarlane R., Munro A., Argyle D.J.;
RT "Isolation and expression of the reverse transcriptase component of
RT the Canis familiaris telomerase ribonucleoprotein (dogTRT)."
RL Gene 336:105-113(2004).
RW -
CC -1- FUNCTION: Telomerase is a ribonucleoprotein enzyme essential for
CC the replication of chromosome termini in most eukaryotes. It
CC elongates telomeres. It is a reverse transcriptase that adds
CC simple sequence repeats to chromosome ends by copying a template
CC sequence within the RNA component of the enzyme.
CC -1- SUBUNIT: Component of the telomerase ribonucleoprotein complex at
CC least composed of TRF1, EST1, POT1 and a telomerase RNA template
CC component (TER). Interacts with PINK1 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the reverse transcriptase family.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL/AF380351; AAC02791.1; -  
 DR PROSITE; PS50876; RT POL; 1.  
 DR DNA-binding; Nuclear protein; Ribonucleoprotein;  
 DR RNA-directed DNA polymerase; Telomere; Transferrase.  
 KM DOMAINE 595 926 Reverse transcriptase.  
 SQ SEQUENCE 1123 AA; 124823 MW; P5F55D791106C1A3 CRC64;

Query Match 78.7%; Score 37; DB 1; Length 1123;  
 Best Local Similarity 66.7%; Pred. No. 1.1e+03;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPGTKT 9  
 Db 283 GPGTKRPT 291

RESULT 12  
 ID Q684M7 PRELIMINARY; PRT; 1184 AA.  
 AC Q684M7;  
 DT 25-OCT-2004 (TREMblrel. 28, Created)  
 DT 25-OCT-2004 (TREMblrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMblrel. 28, Last annotation update)  
 DE Tyrosine kinase 2.  
 GN Name=TYK2;  
 OS Sue scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OC NCBI\_TaxID=9623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98343959; PubMed=9677330;  
 RA Sullivan M., Rens G., Begg F., Gordon L., Olsen A.S., Houslay M.D.;  
 RT "Identification and characterization of the human homologue of  
 the short PDE4A cAMP-specific phosphodiesterase RD1 (PDE4A1) by analysis  
 of the human HSPDE4A gene locus located at chromosome 19p13.2.";  
 RL Blocham. J. 333:693-703 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21203792; PubMed=11306681;  
 RA Rens G., Begg F., Ross A., Mackenzie C., McPhee I., Campbell L.,  
 RA Hutton E., Sullivan M., Houslay M.D.;  
 RT "Molecular cloning, genomic positioning, promoter identification, and  
 characterization of the novel cyclic amp-specific phosphodiesterase  
 PDE4A10.";  
 RL Mol. Pharmacol. 59:996-1011 (2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Leeb T., Martins-Wess F., Kuiper H., Diehl O., Muller M.;  
 RT "Molecular characterization of the porcine TYK2 gene on SSC 2q1.3-  
 q2.1.";  
 RL Cytogenet. Genome Res. 107:103-107 (2004).  
 CC -1- FUNCTION: Tyrosine kinase of the non-receptor type (by  
 similarity).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. JAK  
 CC subfamily.  
 CC EMBL; AJ632303; CAG5148.1; -  
 DR GO; GO:0016301; P:kinase activity; IEA.  
 DR InterPro; IPR000299; Band\_4.1.  
 DR InterPro; IPR009127; JAK.  
 DR InterPro; IPR011009; Kinase like.  
 DR InterPro; IPR009131; Nonreceptor tyrosine.  
 DR InterPro; IPR007119; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001245; Tyr\_kinase.

DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00069; Kinase; 2.  
 DR PRINTS; PR01823; JAKSKINASE.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRINTS; PR01827; YKINASETYK2.  
 DR ProDom; PD000001; Prot\_kinase; 2.  
 DR SMART; SM00295; B41; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR SMART; SM00230; S\_TKc; 2.  
 DR SMART; SM00219; TYRKc; 2.  
 DR PROSITE; PS50057; FERM\_3; 1.  
 DR PROSITE; PS500107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 2.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KM ATP-binding; Kinase; SH2 domain; Transferrase; Tyrosine-protein kinase.  
 SQ SEQUENCE 1184 AA; 123309 MW; 947255CC320B040B CRC64;

Query Match 78.7%; Score 37; DB 2; Length 1184;  
 Best Local Similarity 85.7%; Pred. No. 1.1e+03;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGTKT 7  
 Db 127 GPGTKET 133

RESULT 13  
 ID Q9L470 PRELIMINARY; PRT; 1245 AA.  
 AC Q9L470;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)  
 DE Accumulation-associated protein precursor.  
 GN Name=asp;  
 OS Staphylococcus epidermidis.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OC NCBI\_TaxID=1282;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Pe2A;  
 RA Hussain Shalikh M., Hellmann C., Peters G., Herrmann M.;  
 RT Submitted (S8P-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 CC an amide bond (by similarity).  
 DR EMBL; AJ249487; CAB7251.1; -  
 DR GO; GO:0009986; C:cell surface; IEA.  
 DR GO; GO:0005618; C:cell wall; IEA.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro; IPR008161; C1g\_helix.  
 DR InterPro; IPR008160; Collagen.  
 DR InterPro; IPR004829; Collagen\_antigen.  
 DR InterPro; IPR011098; G5.  
 DR InterPro; IPR005877; Gpos\_YSIRK.  
 DR InterPro; IPR001899; Gram\_pos\_anchor.  
 DR Pfam; PF07501; G5; 4.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF04650; YSIRK\_signal; 1.  
 DR ProDom; PD000007; C1g\_helix; 1.  
 DR ProDom; PD153432; Surface antigen; 2.  
 DR TIGRFAMS; TIGR01167; LPTXG\_anchor; 1.  
 DR TIGRFAMS; TIGR01168; YSIRK\_signal; 1.  
 KM PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 KM Cell wall; Peptidoglycan-anchor; Signal.  
 FT SIGNAL 1 52 Potential.  
 FT CHAIN 53 1245 accumulation-associated protein.  
 SQ SEQUENCE 1245 AA; 131966 MW; B68425AAB45604D3 CRC64;

Query Match 78.7%; Score 37; DB 2; Length 1245;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+03;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GPGTKTXT 9  
 DB 771 GPGTKTIT 779

## RESULT 14

PLS\_STAMP STANDARD; PRT; 1469 AA.

AC Q6CQD9; 05-JUL-2004 (Rel. 44, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Accumulation-associated protein precursor.  
 GN OrderedLocustNames=SR0175;  
 OS Staphylococcus epidermidis  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=1282;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SMPAIN-ATCC 12228;  
 RX PubMed=12850922;  
 RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,  
 RA Qian Z.-Q., Zhao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,  
 RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.,  
 RT "Genome-based analysis of virulence genes in a non-biofilm-forming  
 Staphylococcus epidermidis strain (ATCC 12228).";  
 RL Mol. Microbiol. 49:1577-1593(2003).  
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 CC an amide bond (Potential).  
 CC -1- SIMILARITY: Contains 7 G5 repeats.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB016744; AA003772.1; -  
 DR InterPro: IPR008160; Collagen.  
 DR InterPro: IPR008985; Cona\_like\_1ec\_g1.  
 DR InterPro: IPR004829; Caurface\_antigen.  
 DR InterPro: IPR011098; G5.  
 DR InterPro: IPR005877; Gpos\_Y5IRK.  
 DR InterPro: IPR001899; Gram\_pos\_anchor.  
 DR Pfam; PF07501; G5; 7.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF04650; Y5IRK\_signal; 1.  
 DR ProDom; PD000007; Clg\_helix; 2.  
 DR ProDom; PD153432; Caurface\_antigen; 2.  
 DR TIGRFAMs; TIGR01167; LPTXG\_anchor; 1.  
 DR TIGRFAMs; TIGR01168; Y5IRK\_signal; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 KW Cell wall; Complete proteome; Peptidoglycan-anchor; Repeat; Signal.  
 FT SIGNAL 1 52 Potential  
 FT CHAIN 53 1435 Accumulation-associated protein.  
 FT PROPEP 1436 1469 Removed by sortase (Potential).  
 FT DOMAIN 446 528 G5 1.  
 FT DOMAIN 574 656 G5 2.  
 FT DOMAIN 702 784 G5 3.  
 FT DOMAIN 830 912 G5 4.  
 FT DOMAIN 958 1040 G5 5.  
 FT DOMAIN 1086 1168 G5 6.  
 FT DOMAIN 1214 1296 G5 7.  
 FT SITE 1432 1436 LPTXG sorting signal (Potential).  
 FT MOD\_RES 1435 1435 Pentaglycyl murein peptidoglycan amidated  
 FT threonine (Potential).  
 SQ SEQUENCE 1469 AA; 157137 MW; 9D9333095B680182 CRC64;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GPGTKTXT 9  
 DB 611 GPGTKTIT 619

## RESULT 15

06UV37 PRELIMINARY; PRT; 1477 AA.

AC Q6UV37; 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Accumulation-associated protein.  
 GN OrderedLocustNames=SR0175;  
 OS Staphylococcus epidermidis  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=1282;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Robde H., Mack D.;  
 RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 CC an amide bond (by similarity).  
 CC EMBL; AY359816; AA083699.1; -  
 DR GO; GO:0009986; C:cell surface; IEA.  
 DR GO; GO:0005618; C:cell wall; IEA.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro: IPR008160; Collagen.  
 DR InterPro: IPR004829; Caurface\_antigen.  
 DR InterPro: IPR011098; G5.  
 DR InterPro: IPR005877; Gpos\_Y5IRK.  
 DR InterPro: IPR001899; Gram\_pos\_anchor.  
 DR Pfam; PF07501; G5; 6.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF04650; Y5IRK\_signal; 1.  
 DR ProDom; PD153432; Caurface\_antigen; 2.  
 DR TIGRFAMs; TIGR01167; LPTXG\_anchor; 1.  
 DR TIGRFAMs; TIGR01168; Y5IRK\_signal; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 KW Cell wall; Peptidoglycan-anchor.  
 SQ SEQUENCE 1477 AA; 156790 MW; B3AAD34FE7F37F14 CRC64;

Query Match 78.7%; Score 37; DB 2; Length 1477;  
 Best Local Similarity 77.8%; Pred. No. 1.4e+03;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GPGTKTXT 9  
 DB 741 GPGTKTIT 749

Search completed: April 18, 2005, 08:15:36  
 Job time : 40.5766 secs

Query Match 78.7%; Score 37; DB 1; Length 1469;  
 Best Local Similarity 77.8%; Pred. No. 1.4e+03;



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 07:10:15 ; Search time 69.6934 Seconds  
(without alignments)  
77.692 Million cell updates/sec

Title: US-10-652-334-2  
Perfect score: 57  
Sequence: 1 RILXCASMXAVDXL 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	51	89.5	14	7	ABW01178	Abw01178 Saccharom
2	51	89.5	14	8	ADP44104	ADP44104 Yeast tra
3	51	89.5	15	3	AAV77805	AAV77805 Motif II
4	40.5	71.1	380	7	AAV77814	AAV77814 Yeast Upf
5	40.5	71.1	380	8	ABW01205	ABW01205 Saccharom
6	40.5	71.1	380	7	ADP44131	ADP44131 Yeast hel
7	40.5	71.1	611	8	ADS21255	ADS21255 Bacterial
8	40.5	71.1	828	4	AAW39674	AAW39674 Human pol
9	40.5	71.1	828	4	AAW39277	AAW39277 Human pro
10	40.5	71.1	828	6	ABW58327	ABW58327 NM_015046
11	40.5	71.1	828	8	ADP23318	ADP23318 PRO polyp
12	40.5	71.1	829	4	AAW80279	AAW80279 Human pro
13	40.5	71.1	829	4	AAW41460	AAW41460 Human pol
14	40.5	71.1	966	4	AAW79295	AAW79295 Human pro
15	40.5	71.1	971	3	AAW98057	AAW98057 Yeast Upf
16	40.5	71.1	971	6	ABW53412	ABW53412 Protein s
17	40.5	71.1	971	6	ADK64706	ADK64706 Disease t
18	40.5	71.1	2677	6	ABW58288	ABW58288 BCUI156 p
19	39.5	69.3	181	5	AAO17597	AAO17597 E_gosaypi
20	38.5	67.5	712	8	ADS44332	ADS44332 Bacterial
21	37.5	65.8	13	7	ABW01189	ABW01189 Saccharom
22	37.5	65.8	13	8	ADP44115	ADP44115 Yeast tra
23	36.5	64.0	472	7	AAV77815	AAV77815 Yeast Sen
24	36.5	64.0	472	7	ABW01202	ABW01202 Saccharom
25	36.5	64.0	472	8	ADP44128	ADP44128 Yeast hel

26	36.5	64.0	1417	4	ABB63329	Abb63329 Drosophi
27	36.5	64.0	2000	6	ABW53451	Abw53451 Protein s
28	36.5	64.0	2000	7	ADK64622	ADK64622 Disease t
29	36.5	64.0	2231	8	ADN19177	Adn19177 Bacterial
30	35.5	62.3	685	4	ABB58231	Abb58231 Drosophi
31	35.5	62.3	818	8	ADK68058	ADK68058 Female re
32	35.5	62.3	988	6	AAO26745	AAO26745 988-mer r
33	35.5	62.3	992	8	ADS23757	ADS23757 Bacterial
34	35.5	62.3	1944	8	ADN19868	Adn19868 Bacterial
35	35	61.4	658	8	ADQ59484	Adq59484 Human can
36	34.5	60.5	517	6	ABW9784	ABW9784 Amino aci
37	34.5	60.5	550	6	ABW99781	ABW99781 Amino aci
38	34.5	60.5	804	7	ADB65007	ADB65007 Human pro
39	34.5	60.5	925	8	ADN19434	Adn19434 Bacterial
40	34.5	60.5	935	7	ADJ69900	Adj69900 Human hea
41	34.5	60.5	1002	2	ABW79913	Abw79913 Arabidops
42	34.5	60.5	1043	2	AAW36509	AAW36509 Murine RE
43	34.5	60.5	1118	6	ABG73900	ABG73900 Human REN
44	34.5	60.5	1140	2	AAV21377	AAV21377 Human HUP
45	34.5	60.5	1187	4	AAU07865	AAU07865 Polypepti

## ALIGNMENTS

RESULT 1	ABW01178	standard; peptide: 14 AA.
XX	ABW01178;	
AC	15-JAN-2004.	(first entry)
XX		
DT		
XX		
DB	Saccharomyces cerevisiae motif II peptide.	
XX		
KW	Modulator of translation termination; MTTI; helicase B; antiviral;	
KW	therapy; HCSB; nonsense mutation; yeast.	
XX		
OS	Saccharomyces cerevisiae.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference 4	/label= Unknown
FT		/note= "Xaa may be any amino acid"
FT	Misc-difference 9	/label= Unknown
FT		/note= "Xaa may be any amino acid"
FT	Misc-difference 13	/label= Unknown
FT		/note= "Xaa may be any amino acid"
XX		
PR	22-JUL-1999;	99US-00359268.
XX		
PA	22-JUL-1998;	98US-0093685P.
XX		
XX	(UTNR-) UNIV NEW JERSEY MEDICINE & DENTISTRY.	
XX		
PI	Peltz S, Czaplinski K, Dinman JD;	
XX		
DR	WPI; 2003-810549/76.	
XX		
PT	Identifying an agent that increases nonsense suppression, for antiviral	
PT	therapy, by contacting modulator of translation termination (MTTI) in	
PT	Saccharomyces cerevisiae with a test agent, and detecting specific	
PT	binding to MTTI.	
XX		
PS	Disclosure; Col 41-42; Opp; English.	
XX		
CC	The invention relates to a method of identifying an agent that increases	

CC nonsense suppression, by contacting modulator of translation termination  
CC (MTT1) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.  
CC The method is useful for identifying compositions or agents which  
CC increase nonsense suppression. The invention may also be used for  
CC antiviral therapy and for suppression of pathological nonsense mutations.  
CC The present sequence is Saccharomyces cerevisiae motif peptide  
XX  
SQ Sequence 14 AA;  
Query Match 89.5%; Score 51; DB 7; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.0022;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 RIIXCASNXXAVDXL 14  
Db 1 RIIXCASNXXAVDXL 14  
RESULT 2  
ADP44104  
ID ADP44104 standard; peptide: 14 AA.  
XX  
AC ADP44104;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Yeast translation termination modulation protein motif II.  
XX  
XX gene therapy; translation termination; RNA helicase; MTT1;  
XX frameshift frequency; aberrant transcript degradation;  
XX peptidyl transferase modulation; beta-thalassemia; beta-globin;  
XX Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
XX Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;  
XX Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic fibrosis;  
XX Kidney Stone; Familial hypercholesterolemia; Retinitis pigmentosa;  
XX Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; yeast.  
XX  
OS Saccharomyces cerevisiae.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
XX MISC-difference 4 /note= "Any amino acid"  
XX MISC-difference 9 /note= "Any amino acid"  
XX MISC-difference 13 /note= "Any amino acid"  
XX  
XX US2004115787-A1.  
XX  
XX PD 17-JUN-2004.  
XX  
XX 28-AUG-2003; 2003US-00652334.  
XX  
XX 22-JUL-1998; 98US-0093685P.  
XX 22-JUL-1999; 99US-00359268.  
XX  
XX (PELTZ) PELTZ S.  
XX (CZAP) CZAPLINSKI K.  
XX (DINM) DINMAN J D.  
XX  
XX PELTZ S, Czaplinski K, Dinman JD;  
XX  
XX WPI; 2004-449400/42.  
XX  
XX Identifying a test composition or agent that modulates the efficiency of  
XX translation termination comprises contacting the MTT1 with the test  
XX composition or agent, and determining if the test composition or agent  
XX inhibits the MTT1.  
XX  
XX Claim 34; SEQ ID NO 2; 41pp; English.  
XX  
XX The invention relates to a method of identifying a test composition that

CC modulates the efficiency of translation termination comprising contacting  
CC the RNA helicase MTT1 with a composition or agent under conditions  
CC permitting binding between the MTT1 and the composition, detecting  
CC specific binding of the test composition or agent to the MTT1, and  
CC determining if the test composition or agent inhibits the MTT1. The  
CC composition and methods are useful for modulating the fidelity of  
CC translation termination or for identifying agents that affect the  
CC functional activity of mRNAs by altering frameshift frequency, permit  
CC monitoring of a termination event, promote degradation of aberrant  
CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl  
CC transferase activity during initiation, elongation, termination and mRNA  
CC degradation of translation. The agents, which may be antagonists or  
CC agonists, are useful in screening, diagnostic and therapeutic purposes,  
CC for diseases or conditions resulting from or cause premature translation,  
CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular  
CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,  
CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,  
CC Hirschsprung disease, Cystic fibrosis, Kidney Stones, Familial  
CC hypercholesterolemia, Retinitis pigmentosa, or Neurofibromatosis,  
CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents  
CC the amino acid sequence of the yeast translation termination modulation  
CC protein motif II.  
XX  
XX  
SQ Sequence 14 AA;  
Query Match 89.5%; Score 51; DB 8; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.0022;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 RIIXCASNXXAVDXL 14  
Db 1 RIIXCASNXXAVDXL 14  
RESULT 3  
AA77805  
ID AA77805 standard; peptide: 15 AA.  
XX  
XX AA77805;  
XX  
XX 31-MAY-2000. (first entry)  
XX  
XX Motif II comprised in a gene modulating translation termination.  
XX  
XX Helicase B; HCSB; MTT1; modulator of translation termination; eRF1; eRF3;  
XX eukaryotic release factor; peptidyl transferase; beta-thalassemia;  
XX beta-globin; Duchenne/Becker Muscular Dystrophy; anticancer.  
XX  
XX Unidentified.  
XX  
XX OS  
XX  
XX Key Location/Qualifiers  
XX MISC-difference 1.15 /note= "residues indicated Xaa are unspecified"  
XX  
XX WO200005586-A2.  
XX  
XX PD 03-FEB-2000.  
XX  
XX 22-JUL-1999; 99WO-US016802.  
XX 22-JUL-1998; 98US-00120435.  
XX  
XX (UNNE-) UNIV NEW JERSEY.  
XX  
XX PELTZ S, Czaplinski K, Dinman JD;  
XX  
XX WPI; 2000-171458/15.  
XX  
XX New multiprotein complex which can modulate peptidyl transferase activity  
XX during translation, useful to treat diseases associated with peptidyl  
XX transferase activity e.g. Duchenne/Becker Muscular Dystrophy.  
XX  
XX Claim 34; Page 79; 89pp; English.  
XX  
XX

XX The invention provides a new multiprotein complex which can modulate  
 CC peptidyl transferase activity during translation. The complex comprises  
 CC the gene encoding Helicase B (HCSB, renamed MTT1, for Modulator of  
 CC translation termination) and the conserved proteins known to interact and  
 CC carry out translation termination in eukaryotic cells, peptidyl  
 CC eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to  
 CC modulate peptidyl transferase activity during translation in a cell. It  
 CC can be administered therapeutically combined with a carrier in  
 CC pharmaceutical compositions to treat diseases associated with peptidyl  
 CC transferase activity, especially diseases resulting from a nonsense or  
 CC frameshift mutation e.g. beta-thalassemia, beta-globin, Duchenne/Becker  
 CC Muscular Dystrophy etc. It can be used to identify disease conditions  
 CC involving a defect in the complex, by transfecting cells with encoding  
 CC nucleic acid and determining the proportion of defective complex before  
 CC and after transfection. It is also useful to screen for drugs involved in  
 CC peptidyl transferase activity during translation, inhibiting the  
 CC interaction between MTT1 and eRF3 or involved in enhancing translation  
 CC termination. Vectors comprising polynucleotides encoding the complex (or  
 CC antisense sequences) can be constructed and introduced into cells to  
 CC interfere with complex expression and so modulate the efficiency of  
 CC translation termination of mRNA and/or degradation of aberrant  
 CC transcripts in a cell. Agents binding to the complex can be identified  
 CC and included in therapeutic compositions useful as above, and/or used to  
 CC modulate peptidyl transferase activity during translation in cells. They  
 CC are also useful to modulate the efficiency of translation termination of  
 CC mRNA at a nonsense codon and/or promote degradation of aberrant  
 CC transcripts in cells. The method can be used to identify agents/  
 CC compositions modulating binding to MTT1, useful to identify genes.  
 CC Sequences AA177804-812 represent motifs 1-ix comprised in the genes of  
 CC interest, used for modulating translation termination  
 XX

## SQ Sequence 15 AA;

Query Match 89.5%; Score 51; DB 3; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.0024;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIIXCASNXYVDL 14  
 |||||  
 DB 1 RIIXCASNXYVDL 14

## RESULT 4

AA177814  
 ID AA177814 standard; peptide: 380 AA.

XX AC AA177814;

XX DT 31-MAY-2000 (first entry)

XX DE Yeast Upfi protein fragment.

XX KM Helicase B; HCSB; MTT1; modulator of translation termination; eRF1, eRF3;

XX KM eukaryotic release factor; peptidyl transferase; beta-thalassemia;

XX KM beta-globin; Duchenne/Becker Muscular Dystrophy; antianemic; yeast;

XX KM helicase; Upfi.

XX OS Saccharomyces cerevisiae.

XX PN WO200005586-A2.

XX PD 03-FEB-2000.

XX PF 22-JUL-1999; 99WO-US016802.

XX PR 22-JUL-1998; 98US-00120435.

XX PA (UTNE-) UNIV NEW JERSEY.

XX PI Pelcz S, Czaplinski K, Dinman JD;

XX DR WPI; 2000-171456/15.

XX New multiprotein complex which can modulate peptidyl transferase activity  
 PT during translation, useful to treat diseases associated with peptidyl  
 PT transferase activity e.g. Duchenne/Becker Muscular Dystrophy.

PS Example 1; Fig 1; 89p; English.

XX The invention provides a new multiprotein complex which can modulate  
 CC peptidyl transferase activity during translation. The complex comprises  
 CC the gene encoding Helicase B (HCSB, renamed MTT1, for Modulator of  
 CC translation termination) and the conserved proteins known to interact and  
 CC carry out translation termination in eukaryotic cells, peptidyl  
 CC eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to  
 CC modulate peptidyl transferase activity during translation in a cell. It  
 CC can be administered therapeutically combined with a carrier in  
 CC pharmaceutical compositions to treat diseases associated with peptidyl  
 CC transferase activity, especially diseases resulting from a nonsense or  
 CC frameshift mutation e.g. beta-thalassemia, beta-globin, Duchenne/Becker  
 CC Muscular Dystrophy etc. It can be used to identify disease conditions  
 CC involving a defect in the complex, by transfecting cells with encoding  
 CC nucleic acid and determining the proportion of defective complex before  
 CC and after transfection. It is also useful to screen for drugs involved in  
 CC peptidyl transferase activity during translation, inhibiting the  
 CC interaction between MTT1 and eRF3 or involved in enhancing translation  
 CC termination. Vectors comprising polynucleotides encoding the complex (or  
 CC antisense sequences) can be constructed and introduced into cells to  
 CC interfere with complex expression and so modulate the efficiency of  
 CC translation termination of mRNA and/or degradation of aberrant  
 CC transcripts in a cell. Agents binding to the complex can be identified  
 CC and included in therapeutic compositions useful as above, and/or used to  
 CC modulate peptidyl transferase activity during translation in cells. They  
 CC are also useful to modulate the efficiency of translation termination of  
 CC mRNA at a nonsense codon and/or promote degradation of aberrant  
 CC transcripts in cells. The method can be used to identify agents/  
 CC compositions modulating binding to MTT1, useful to identify genes.  
 CC Sequences AA177813-817 represent protein fragments from yeast superfamily  
 CC group I helicases  
 XX

## SQ Sequence 380 AA;

Query Match 71.1%; Score 40.5; DB 3; Length 380;  
 Best Local Similarity 73.3%; Pred. No. 10;  
 Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 RIIXCA-SNXYVDL 14  
 |||||  
 DB 31 RIIXCAPSNVAVDHL 45

## RESULT 5

ABW01205  
 ID ABW01205 standard; protein: 380 AA.

XX AC ABW01205;

XX DT 15-JAN-2004 (first entry)

XX DB Saccharomyces cerevisiae UPFI protein.

XX KM Modulator of translation termination; MTT1, helicase B; antiviral;

XX KM therapy; HCSB; nonsense mutation; yeast; RENT1; HUPFI; UPFI.

XX OS Saccharomyces cerevisiae.

XX PN US6630294-B1.

XX PD 07-OCT-2003.

XX PF 22-JUL-1999; 99US-00359268.

XX PR 22-JUL-1998; 98US-0093685P.

XX PA (UTNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

XX PI Peltz S, Czaplinski K, Dimman JD;  
 XX DR WPI; 2003-810549/76.  
 XX PT Identifying an agent that increases nonsense suppression, for antiviral  
 PT therapy, by contacting modulator of translation termination (MTT1) in  
 PT Saccharomyces cerevisiae with a test agent, and detecting specific  
 PT binding to MTT1.  
 XX PS Disclosure; Col 57-60; 0pp; English.  
 XX CC The invention relates to a method of identifying an agent that increases  
 CC nonsense suppression, by contacting modulator of translation termination  
 CC (MTT1) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.  
 CC The method is useful for identifying compositions or agents which  
 CC increase nonsense suppression. The invention may also be used for  
 CC antiviral therapy and for suppression of pathological nonsense mutations.  
 CC The present sequence is Saccharomyces cerevisiae UPF1 (otherwise termed  
 CC RMT1, HUPF1) protein  
 XX CC  
 XX SQ Sequence 380 AA;  
 QY Query Match 71.1%; Score 40.5; DB 7; Length 380;  
 Db Best Local Similarity 73.3%; Pred. No. 10;  
 Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 1 RILXCA-SNXAVDXL 14  
 31 RILVCAPSNVAVDHL 45  
 RESULT 6  
 ADP4131  
 ID ADP4131 standard; protein; 380 AA.  
 XX AC ADP4131;  
 XX DT 18-NOV-2004 (first entry)  
 XX DE Yeast helicase Upf1.  
 XX DE gene therapy; translation termination; RNA helicase; MTT1;  
 KW frameshift frequency; aberrant transcript degradation;  
 KW peptidyl transferase modulation; beta-thalassemia; beta-globin;  
 KW Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
 KW Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;  
 KW Ovarian Cancer; Wilms Tumour; Hirschprung disease; Cystic fibrosis;  
 KW Kidney Stone; Familial hypercholesterolemia; Retinitis pigmentosa;  
 KW Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; yeast; enzyme.  
 XX OS Saccharomyces cerevisiae.  
 XX PN US2004115787-A1.  
 XX PD 17-JUN-2004.  
 XX PF 28-AUG-2003; 2003US-00652334.  
 XX PR 22-JUL-1998; 98US-0093685P.  
 XX PR 22-JUL-1999; 99US-00359268.  
 XX PA (PELTZ) PELTZ S.  
 XX PA (CZAP/) CZAPLINSKI K.  
 XX PA (DINM/) DINMAN J D.  
 XX PI Peltz S, Czaplinski K, Dimman JD;  
 XX DR WPI; 2004-449400/42.  
 XX PT Identifying a test composition or agent that modulates the efficiency of  
 PT translation termination comprises contacting the MTT1 with the test  
 PT composition or agent, and determining if the test composition or agent

PT inhibits the MTT1.  
 XX PS Disclosure; SEQ ID NO 29; 41pp; English.  
 XX CC The invention relates to a method of identifying a test composition that  
 CC modules the efficiency of translation termination comprising contacting  
 CC the RNA helicase MTT1 with a composition or agent under conditions  
 CC permitting binding between the MTT1 and the composition, detecting  
 CC specific binding of the test composition or agent to the MTT1, and  
 CC determining if the test composition or agent inhibits the MTT1. The  
 CC composition and methods are useful for modulating the fidelity of  
 CC translation termination or for identifying agents that affect the  
 CC functional activity of mRNAs by altering frameshift frequency, permit  
 CC monitoring of a termination event, promote degradation of aberrant  
 CC transcripts, and provide modulators (inhibitors/activators) of peptidyl  
 CC transferase activity during initiation, elongation, termination and mRNA  
 CC degradation of translation. The agents, which may be antagonists or  
 CC agonists, are useful in screening, diagnostic and therapeutic purposes,  
 CC for diseases or conditions resulting from or cause premature translation,  
 CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular  
 CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,  
 CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,  
 CC Hirschprung disease, Cystic fibrosis, Kidney Stones, Familial  
 CC hypercholesterolemia, Retinitis pigmentosa, or Neurofibromatosis,  
 CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents  
 CC the amino acid sequence of the yeast helicase Upf1.  
 XX CC  
 XX SQ Sequence 380 AA;  
 QY Query Match 71.1%; Score 40.5; DB 8; Length 380;  
 Db Best Local Similarity 73.3%; Pred. No. 10;  
 Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 1 RILXCA-SNXAVDXL 14  
 31 RILVCAPSNVAVDHL 45  
 RESULT 7  
 ADS21255  
 ID ADS21255 standard; protein; 611 AA.  
 XX AC ADS21255;  
 XX DT 02-DEC-2004 (first entry)  
 XX DE Bacterial polypeptide #10788.  
 XX KW Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.  
 XX OS Bacteria.  
 XX PN US2003233675-A1.  
 XX PD 18-DEC-2003.  
 XX PF 20-FEB-2003; 2003US-00369493.  
 XX PR 21-FEB-2002; 2002US-0360039P.  
 XX PA (CAOY/) CAO Y.  
 XX PA (HINKLE/) HINKLE G J.  
 XX PA (SLATER/) SLATER S C.  
 XX PA (CHEN/) CHEN X.  
 XX PA (GOLD/) GOLDMAN B S.  
 XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
 XX DR

```

XX  WPI; 2004-061375/06.
DR
XX
XX  New recombinant DNA construct comprising a promoter positioned to provide
PT  for expression of a polynucleotide encoding a polypeptide from a
PT  microbial source, useful for producing plants with improved properties.
PS  Claim 1; SEQ ID NO 10288; 122pp; English.
XX
XX  The invention relates to a recombinant DNA construct comprising a
CC  promoter functional in a plant cell, where the promoter is positioned to
CC  provide for expression of a polynucleotide encoding a polypeptide from a
CC  microbial source. The invention also relates to a transformed plant
CC  comprising the recombinant DNA construct and a method of producing a
CC  transformed plant having an improved property. The plant is a crop plant
CC  such as maize or soybean. The method of producing a transformed plant
CC  having an improved property comprises transforming a plant with the
CC  recombinant DNA construct and growing the transformed plant, where the
CC  polynucleotide or polypeptide is useful for improving plant properties.
CC  The recombinant DNA construct is useful for producing plants with
CC  improved plant properties, e.g. improved cold, heat or drought tolerance,
CC  tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC  increased resistance to plant disease, better growth rate by modification
CC  of the cell cycle pathway with plant growth regulators, increased rate of
CC  homologous recombination, modified seed oil or protein yield and/or
CC  content, improved yield by modification of carbohydrate, nitrogen or
CC  phosphorus use and/or uptake, by modification of photosynthesis or by
CC  providing improved plant growth and development under at least one stress
CC  condition, improved lignin production or improved galactomannan
CC  production. This sequence represents a bacterial polypeptide used in the
CC  scope of the invention. Note: The sequence data for this patent did not
CC  form part of the printed specification but was obtained in electronic
CC  format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ  Sequence 611 AA;
XX
XX  Query Match          71.1%; Score 40.5; DB 8; Length 611;
XX  Best Local Similarity 66.7%; Pred. No. 17;
XX  Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
XX
OY  1 RILKCA-SNXAVDXL 14
XX  |||||
XX  215 RVLVCAPSNAAVDVL 229
DB
XX
RESULT 8
AAM39674
ID  AAM39674 standard; protein; 828 AA.
XX
AC  AAM39674;
XX
DT  22-OCT-2001 (first entry)
XX
XX  Human polypeptide SEQ ID NO 2819.
XX
XX  Human; nocrotic; immunosuppressant; cytostatic; gene therapy; cancer;
XX  peripheral nervous system; neuropathy; central nervous system; CNS;
XX  Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX  amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX  chemokine; chromolytic; drug screening; arthritis; inflammation;
XX  leukaemia.
XX
OS  Homo sapiens.
XX
XX  WO200153312-A1.
XX
XX  26-JUL-2001.
XX
XX  26-DEC-2000; 2000WO-US034263.
XX
XX  23-DEC-1999; 99US-00471275.
XX  21-JAN-2000; 2000US-00488725.
XX  25-APR-2000; 2000US-00552317.

```

```

PR  20-JUN-2000; 2000US-00598042.
PR  19-JUL-2000; 2000US-00620312.
PR  03-AUG-2000; 2000US-00653450.
PR  14-SEP-2000; 2000US-00662191.
PR  19-OCT-2000; 2000US-00693036.
PR  29-NOV-2000; 2000US-00727344.
XX
XX  (HYSR-) HYSEQ INC.
XX
XX  Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX  Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX  Zhou P, Goodrich R, Dimanac RT;
XX
XX  WPI; 2001-442253/47.
XX  N-PSDB; AA158830.
XX
XX  Novel nucleic acids and polypeptides, useful for treating disorders such
XX  as central nervous system injuries.
XX
XX  Example 4; SEQ ID NO 2819; 10078pp; English.
XX
XX  The invention relates to human nucleic acids (AA157798-AA161369) and the
XX  encoded polypeptides (AAM38642-AA42213) with nocrotic,
XX  immunosuppressant and cytostatic activity. The polynucleotides are useful
XX  in gene therapy. A composition containing a polypeptide or polynucleotide
XX  of the invention may be used to treat diseases of the peripheral nervous
XX  system, such as peripheral nervous injuries, peripheral neuropathy and
XX  localized neuropathies and central nervous system diseases, such as
XX  Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX  lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX  utilisation of the activities such as: Immune system suppression,
XX  Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX  and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX  assays for receptor activity, arthritis and inflammation, leukaemia and
XX  C.N.S disorders. Note: The sequence data for this patent did not form
XX  part of the printed specification
XX
SQ  Sequence 828 AA;
XX
XX  Query Match          71.1%; Score 40.5; DB 4; Length 828;
XX  Best Local Similarity 66.7%; Pred. No. 24;
XX  Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
XX
OY  1 RILKCA-SNXAVDXL 14
XX  |||||
XX  153 RVLVCAPSNAAVDVL 167
DB
XX
RESULT 9
AAB92977
ID  AAB92977 standard; protein; 828 AA.
XX
AC  AAB92977;
XX
DT  26-JUN-2001 (first entry)
XX
XX  Human protein sequence SEQ ID NO:11690.
XX
XX  Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX  Homo sapiens.
XX
XX  EP1074617-A2.
XX
XX  07-FEB-2001.
XX
XX  28-JUL-2000; 2000EP-00116126.
XX
XX  29-JUL-1999; 99JP-00248026.
XX  27-AUG-1999; 99JP-00300253.
XX  11-JAN-2000; 2000JP-00118776.
XX  02-MAY-2000; 2000JP-00183767.
XX  09-JUN-2000; 2000JP-00241899.

```

XX	(HELI-) HELIX RES INST.
PA	Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
P1	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR	WPI; 2001-318749/34.
PT	Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PR	length cDNAs defined in the specification, and for the detection and/or
PT	diagnosis of the abnormality of the proteins encoded by the full-length
XX	cDNAs.
PS	Claim 8; SEQ ID NO 11690; 2537bp + Sequence listing; English.
XX	The present invention describes primer sets for synthesizing 5602 full-
CC	length cDNAs defined in the specification. Where a primer set comprises:
CC	(a) an oligo-dT primer and an oligonucleotide which complementary to the
CC	complementary strand of a polynucleotide which comprises one of the 5602
CC	nucleotide sequences defined in the specification; where the
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC	of an oligonucleotide comprising a sequence complementary to the
CC	complementary strand of a polynucleotide which comprises a 5'-end
CC	sequence and an oligonucleotide comprising a sequence complementary to a
CC	polynucleotide which comprises a 3'-end sequence, where the
CC	oligonucleotide comprises at least 15 nucleotides and the combination of
CC	the 5'-end sequence/3'-end sequence is selected from those defined in the
CC	specification. The primer sets can be used in antisense therapy and in
CC	gene therapy. The primers are useful for synthesising polynucleotides,
CC	particularly full-length cDNAs. The primers are also useful for the
CC	detection and/or diagnosis of the abnormality of the proteins encoded by
CC	the full-length cDNAs. The primers allow obtaining of the full-length
CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC	represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC	oligonucleotides, all of which are used in the exemplification of the
CC	present invention
XX	
SQ	Sequence 828 AA;
Query Match	71.1%; Score 40.5; DB 4; Length 828;
Best Local Similarity	66.7%; Pred. No. 24;
Matches	10; Conservative 1; Mismatches 3; Indels 1; Gaps 1.
OY	1 RILKCA-SNXAVDXL 14 
DB	153 RVLWCAPSNAVADEL 167 
RESULT 10	
ABR58327	
ID	ABR58327 standard; protein; 828 AA.
XX	
AC	ABR58327;
XX	
DT	08-JUL-2003 (first entry)
DE	NM_015046 protein #SEQ ID 96.
XX	
KM	Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW	drug discovery; clinical medicine; forensic medicine.
XX	
OS	Homo sapiens.
XX	
PN	WO2003029421-A2.
PD	10-APR-2003.
PF	02-OCT-2002; 2002MO-US031287.
XX	
PR	03-OCT-2001; 2001US-0326526F.
PR	14-MAY-2002; 2002US-00141194.
XX	

```

PA (ORIG-)ORIGENE TECHNOLOGIES INC.
XX
XX Sun Z, Li X, Fan W, Kovacs KF, Jay G;
PI
XX WPI; 2003-381623/36.
XX
XX New isolated human differentially-regulated breast cancer polynucleotide
XX and polypeptide, useful for diagnosing, staging, prognosticating,
XX preventing and/or treating diseases and conditions relating to breast
XX cancer.
XX
XX Disclosure; SEQ ID NO 96; 127bp + Sequence Listing; English.
XX
XX The invention relates to isolated polynucleotides which are
XX differentially-regulated in breast cancer. The methods and compositions
XX of the present invention are useful for detecting, diagnosing, staging,
XX monitoring, prognosticating, preventing and/or treating diseases and
XX conditions relating to breast cancer, and may be used in gene therapy or
XX antisense therapy. They can also be used in research, drug discovery,
XX clinical medicine and forensic medicine. Sequences given in records
XX ABR8834-ABR58345 represent known polypeptides encoded by polynucleotides
XX that are differentially-regulated in breast cancer. NOTE: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 828 AA:
XX
XX Query Match 71.1%; Score 40.5; DB 6; Length 828;
XX Best Local Similarity 66.7%; Pred. No. 24;
XX Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
XX
XX 1 RILXCA-SNAXVXL 14
XX | : | | | | | | |
XX
XX DB 153 RVLWCAPSNAVDL 167
XX
XX RESULT 11
XX ADP23318
XX ID ADP23318 standard; protein; 828 AA.
XX
XX AC ADP23318;
XX
XX DT 18-NOV-2004 (first entry)
XX
XX DE PRO polypeptide SEQ ID NO:412.
XX
XX PRO; antiinflammatory; antiarthritic; antineumatic; immunosuppressive;
XX osteoporitic; antidiabetic; dermatological; antiproliferic; antiallergic;
XX antiaesthetic; hepatotropic; respiratory; gene therapy; immune system.
XX
XX Unidentified.
XX
XX OS
XX
XX PN WO2004041170-A2.
XX
XX PD 21-MAY-2004.
XX
XX PF 30-OCT-2003; 2003WO-US034312.
XX
XX PR 01-NOV-2002; 2002US-0423394P.
XX
XX PA (GETH ) GENENTECH INC.
XX
XX PI Clark H, Schoenfeld J, Van Lookeren W, Williams PM, Wood WI;
XX Wu TD;
XX
XX WPI; 2004-419628/39.
XX
XX DR N-PSDB; ADP23317.
XX
XX New PRO polypeptides and polynucleotides, useful for treating e.g.
XX erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
XX renal disease, or demyelinating diseases of the central or peripheral
XX nervous system.

```

XX Claim 7; SEQ ID NO 412; 2940bp; English.  
PS  
XX The invention relates to a novel isolated nucleic acid and the PRO  
CC polypeptide encoded by it. A protein of the invention has  
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
CC antitachymatic, hepatotropic, and respiratory activity. A polynucleotide  
CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
CC agonist, antagonist, or antibody that specifically binds to the  
CC polypeptide is useful for treating an immune related disorder such as  
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
CC disease, a demyelinating disease of the central or peripheral nervous  
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,  
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary  
CC disease, infectious or autoimmune chronic active hepatitis, primary  
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
CC disease, asthma, allergic rhinitis, atopic dermatitis, food  
CC hypersensitivity, urticaria, an immunologic disease of the lung,  
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
CC pneumonitis, a transplantation associated disease, graft rejection or  
CC graft-versus-host disease. The present sequence represents a PRO protein  
CC of the invention.  
XX  
SQ Sequence 828 AA;  
Query Match 71.1%; Score 40.5; DB 8; Length 828;  
Best Local Similarity 66.7%; Pred. No. 24;  
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
QY 1 RILXCA-SNXAVDXL 14  
Db 153 RVLVCAPSNAVDEL 167  
RESULT 12  
AAM80279  
ID AAM80279 standard; protein; 829 AA.  
AC AAM80279;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human protein SEQ ID NO 3925:  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.  
XX  
OS Homo sapiens.  
XX  
PN WO200157190-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 05-FEB-2001; 2001WO-US004098.  
XX  
XX 03-FEB-2000; 2000US-00496914.  
PR 27-APR-2000; 2000US-00560875.  
PR 20-JUN-2000; 2000US-00598075.  
PR 19-JUL-2000; 2000US-00620325.  
PR 01-SEP-2000; 2000US-00654936.  
PR 15-SEP-2000; 2000US-00663561.  
PR 20-OCT-2000; 2000US-00693325.  
PR 30-NOV-2000; 2000US-00728422.

XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YF, Liu C, Demanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
PI Ma Y, Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;  
DR WPI/ 2001-476283/51.  
DR N-PSDB; AAK53412.  
XX  
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful  
PT in diagnosis and gene therapy.  
XX  
PS Claim 20; Page 469-470; 6221pp; English.  
XX  
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
CC sequence listing were missing at the time of publication  
XX  
SQ Sequence 829 AA;  
Query Match 71.1%; Score 40.5; DB 4; Length 829;  
Best Local Similarity 66.7%; Pred. No. 24;  
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
QY 1 RILXCA-SNXAVDXL 14  
Db 154 RVLVCAPSNAVDEL 168  
RESULT 13  
AAM41460  
ID AAM41460 standard; protein; 829 AA.  
AC AAM41460;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 6391.  
XX  
KW Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US034263.  
XX  
XX 23-DEC-1999; 99US-00471275.  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
PR 20-JUN-2000; 2000US-00598042.  
PR 19-JUL-2000; 2000US-00620312.  
PR 03-AUG-2000; 2000US-00653450.  
PR 14-SEP-2000; 2000US-00662191.  
PR 19-OCT-2000; 2000US-00693036.



```

PR 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Aaundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Dermanac RT;
XX
XX MPI: 2001-442253/47.
DR N-PSDB; AA160616.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Example 2; SEQ ID NO 6391; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AA38642-AA44213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
XX SQ Sequence 829 AA;
XX
XX Query Match 71.1%; Score 40.5; DB 4; Length 829;
XX Best Local Similarity 66.7%; Pred. No. 24;
XX Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
XX
XX 1 RILXCA-SNXAVDXL 14
XX |:| | | | | | | | | |
DB 154 RVLVCAPSNAVDEL 168

```

```

PR 30-NOV-2000; 2000US-00728422.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Dermanac RT, Aaundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX MPI: 2001-476283/51.
DR N-PSDB; AA52428.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
XX Claim 20; Page 4368-4370; 6221bp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAW8323-AAW80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAW80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
XX SQ Sequence 966 AA;
XX
XX Query Match 71.1%; Score 40.5; DB 4; Length 966;
XX Best Local Similarity 66.7%; Pred. No. 28;
XX Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
XX
XX 1 RILXCA-SNXAVDXL 14
XX |:| | | | | | | | | |
DB 291 RVLVCAPSNAVDEL 305

```

```

RESULT 14
AAW79295
ID AAW79295 standard; protein; 966 AA.
XX
XX AAW79295;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human protein SEQ ID NO 1957.
DE
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorder; arthritis; inflammation.
XX
XX Homo sapiens.
OS
XX
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US004098.
XX
XX 03-FEB-2000; 2000US-00496914.
XX 27-APR-2000; 2000US-00560875.
XX 20-JUN-2000; 2000US-00598075.
XX 19-JUL-2000; 2000US-00620325.
XX 01-SEP-2000; 2000US-00654936.
XX 15-SEP-2000; 2000US-00663561.
XX 20-OCT-2000; 2000US-00693325.

```

```

RESULT 15
AAV98057
ID AAV98057 standard; protein; 971 AA.
XX
XX AAV98057;
AC
XX
XX 07-SEP-2000 (first entry)
DT
XX
XX Yeast Upf1p.
DE
XX
XX Yeast; nonsense-mediated mRNA decay pathway; NMD2; UPF3; UPF1; upf1p;
XX upf1p; gene therapy; beta-thalassemia; cancer;
XX polycystic kidney disease; Duchenne muscular dystrophy; Fanconi anaemia;
XX haemophilia; hypercholesterolemia; neurofibromatosis; Tay-Sachs disease;
XX glycogen storage disease; cystic fibrosis; adenomatous polyposis coli;
XX Cowden disease; Maple syrup urine disease; Wilson disease;
XX Niemann-Pick disease; Turcot syndrome; McArdle disease;
XX ornithine transcarbamylase deficiency.
XX
XX Saccharomyces cerevisiae.
OS
XX
XX US6071700-A.
XX
XX 06-JUN-2000.
XX
XX 21-OCT-1998; 98US-00177431.
XX
XX 20-JAN-1995; 95US-00375300.
XX 21-OCT-1997; 97US-00955472.
XX
XX (UTWA-) UNIV MASSACHUSETTS.
XX

```



PI Jacobson AS, He F;  
 XX MPI; 2000-422078/36.  
 DR N-PSDB; AAA39454.  
 XX

PT Testing for compounds able to modulate the nonsense mediated mRNA decay  
 PT pathway for increasing endogenous protein production by incubating a  
 PT candidate compound with a cell having a gene with a nonsense mutation and  
 PT a gene without mutation.  
 XX

PS Disclosure: Fig 6; 57pp; English.  
 XX

CC The nonsense-mediated mRNA decay pathway has an important role in the  
 CC degradation of mRNA transcripts that contain a stop codon. The products  
 CC of the UPB1 and UPB3 genes (proteins Upf1p and Upf3p, respectively) are  
 CC essential components of this pathway. Mutations in these genes or  
 CC inhibition of Upf1p or Upf3p function would lead to stabilisation of mRNA  
 CC containing premature stop codons. The present sequence is the UPB1  
 CC protein, Upf1p, from *Saccharomyces cerevisiae*. The carboxy terminal of  
 CC the NMD2 gene protein, Nmd2p, has been found to bind to Upf1p, leading to  
 CC inhibition of the nonsense-mediated mRNA decay pathway. Inhibition of  
 CC this pathway is a useful means of treating disorders caused by the  
 CC presence of nonsense mutations, e.g. breast cancer, polycystic kidney  
 CC diseases, Niemann-Pick disease, adenomatous polyposis coli, cystic  
 CC fibrosis, Fanconi's anaemia, haemophilia, hypercholesterolaemia,  
 CC neurofibromatosis, ornithine transcarbamylase deficiency, retinoblastoma,  
 CC glycogen storage disease, McArdle disease, polycystic kidney disease,  
 CC disease, Cowden disease, Wilson disease, colorectal cancer, Tay-Sachs  
 CC disease, muscular dystrophy, adrenocortical carcinoma, Li-Fraumeni  
 CC syndrome, Maple syrup urine disease, Turcot syndrome or beta-thalassaemia  
 XX

SQ Sequence 971 AA;

Query Match 71.1%; Score 40.5; DB 3; Length 971;  
 Best Local Similarity 73.3%; Pred. No. 28;  
 Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 Oy 1 RILKCA-SNXAVDXL 14  
 ||| ||| ||| |||  
 Db 454 RILVCAPSNVAVDHL 468

Search completed: April 18, 2005, 08:03:43  
 Job time : 72.8045 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using BW model

Run on: April 18, 2005, 07:37:21 ; Search time 17.4745 Seconds  
(without alignments)  
59.807 Million cell updates/sec

Title: US-10-652-334-2  
Sequence: 1 RILKCSNKAVDXL 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	89.5	14	US-09-359-268A-2	Sequence 2, Appl
2	40.5	71.1	380	US-09-359-268A-29	Sequence 29, Appl
3	40.5	71.1	971	US-08-724-354D-22	Sequence 22, Appl
4	40.5	71.1	971	US-09-270-984A-22	Sequence 22, Appl
5	40.5	71.1	971	US-09-177-431-8	Sequence 8, Appl
6	37.5	65.8	13	US-09-359-268A-13	Sequence 13, Appl
7	37.5	65.8	197	US-09-248-796A-18135	Sequence 18135, A
8	36.5	64.0	472	US-09-359-268A-26	Sequence 26, Appl
9	35.5	62.3	993	US-09-538-092-1100	Sequence 1100, Ap
10	34.5	60.5	1043	US-08-724-354D-4	Sequence 4, Appl
11	34.5	60.5	1043	US-09-270-984A-4	Sequence 4, Appl
12	34.5	60.5	1118	US-08-724-354D-2	Sequence 2, Appl
13	34.5	60.5	1118	US-09-270-984A-2	Sequence 2, Appl
14	34.5	60.5	1140	US-09-949-016-10116	Sequence 10116, A
15	34	59.6	259	US-09-252-991A-16855	Sequence 16855, A
16	33.5	58.8	242	US-09-248-796A-19108	Sequence 19108, A
17	32.5	57.0	414	US-09-358-268A-28	Sequence 28, Appl
18	32.5	57.0	414	US-09-538-092-483	Sequence 483, App
19	32	56.1	270	US-09-248-796A-20535	Sequence 20535, A
20	32	56.1	387	US-09-491-577-60	Sequence 60, Appl
21	32	56.1	713	US-09-198-452A-640	Sequence 640, App
22	32	56.1	714	US-09-438-185A-601	Sequence 601, App
23	31.5	55.3	519	US-09-248-796A-15238	Sequence 15238, A
24	31	54.4	82	US-09-489-039A-10394	Sequence 10394, A
25	31	54.4	116	US-09-489-039A-7198	Sequence 7198, Ap
26	31	54.4	150	US-09-270-767-45608	Sequence 45608, A
27	31	54.4	222	US-09-328-352-6740	Sequence 6740, Ap

28	31	54.4	271	US-09-252-991A-26683	Sequence 26683, A
29	31	54.4	294	US-09-408-020-12	Sequence 12, Appl
30	31	54.4	295	US-09-252-991A-26676	Sequence 26676, A
31	31	54.4	323	US-09-902-540-10675	Sequence 10675, A
32	31	54.4	345	US-09-248-796A-20118	Sequence 20118, A
33	31	54.4	358	US-09-252-991A-28728	Sequence 28728, A
34	31	54.4	362	US-08-846-762-14	Sequence 14, Appl
35	31	54.4	362	US-08-846-762-83	Sequence 83, Appl
36	31	54.4	510	US-09-538-092-603	Sequence 603, App
37	31	54.4	965	US-09-538-092-650	Sequence 650, App
38	31	54.4	2237	US-08-455-543A-48	Sequence 48, Appl
39	31	54.4	2237	US-08-223-305C-48	Sequence 48, Appl
40	31	54.4	2237	US-09-268-163-8	Sequence 8, Appl
41	31	54.4	2337	US-08-713-118-2	Sequence 2, Appl
42	31	54.4	2337	US-09-452-007-2	Sequence 2, Appl
43	31	54.4	2339	US-08-455-543A-47	Sequence 47, Appl
44	31	54.4	2339	US-08-223-305C-47	Sequence 47, Appl
45	31	54.4	2339	US-09-268-163-6	Sequence 6, Appl

## ALIGNMENTS

```

RESULT 1
US-09-359-268A-2
: Sequence 2, Application US/09359268A
: Patent No. 6630294
: GENERAL INFORMATION:
: APPLICANT: Peltz, Stuart
: APPLICANT: Czaplinski, Kevin
: TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
: TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USBS
: FILE REFERENCE: 601-1-85N
: CURRENT APPLICATION NUMBER: US/09/359,268A
: PRIOR FILING DATE: 1999-07-22
: PRIOR APPLICATION NUMBER: 60/093,685
: NUMBER OF SEQ ID NOS: 32
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 2
: LENGTH: 14
: TYPE: PRT
: ORGANISM: Saccharomyces cerevisiae
: OTHER INFORMATION: Xaa = any amino acid
US-09-359-268A-2

Query Match      89.5% Score 51; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00038; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0;

Cy      1 RILKCSNKAVDXL 14
Db      1 RILKCSNKAVDXL 14

RESULT 2
US-09-359-268A-29
: Sequence 29, Application US/09359268A
: Patent No. 6630294
: GENERAL INFORMATION:
: APPLICANT: Peltz, Stuart
: APPLICANT: Czaplinski, Kevin
: TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
: TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USBS
: FILE REFERENCE: 601-1-85N
: CURRENT APPLICATION NUMBER: US/09/359,268A
: CURRENT FILING DATE: 1999-07-22
: PRIOR APPLICATION NUMBER: 60/093,685

```





FEATURE:  
NAME/KEY: misc feature  
LOCATION: (0)-(0)  
OTHER INFORMATION: Polypeptide Accession Number P38935  
US-09-538-092-1100

Query Match 62.3%; Score 35.5; DB 4; Length 993;  
Best Local Similarity 60.0%; Pred. No. 74;  
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

OY 1 ILKCA-SNXAVDXL 14  
DB 237 VLVCAPSNTAVDNL 251

RESULT 10  
US-08-724-354D-4  
Sequence 4, Application US/08724354D  
Patent No. 5994119  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,354D  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/016,482  
FILING DATE: 29-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Halle, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1043 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-724-354D-4

Query Match 60.5%; Score 34.5; DB 2; Length 1043;  
Best Local Similarity 64.3%; Pred. No. 1,2e+02;  
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

OY 2 ILKCA-SNXAVDXL 14  
DB 440 VLVCAPSNTAVDNL 453

RESULT 11  
US-09-270-984A-4  
Sequence 4, Application US/09270984A  
Patent No. 6048965  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF

TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/270,984A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/724,354  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Halle, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1043 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-270-984A-4

Query Match 60.5%; Score 34.5; DB 3; Length 1043;  
Best Local Similarity 64.3%; Pred. No. 1,2e+02;  
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

OY 2 ILKCA-SNXAVDXL 14  
DB 440 VLVCAPSNTAVDNL 453

RESULT 12  
US-08-724-354D-2  
Sequence 2, Application US/08724354D  
Patent No. 5994119  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,354D  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/016,482  
FILING DATE: 29-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Halle, Lisa A.

REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1118 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-724-354D-2

Query Match 60.5%; Score 34.5; DB 2; Length 1118;  
Best Local Similarity 64.3%; Pred. No. 1.4e+02;  
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 2 ILXCA-SNXAVDXL 14  
:|||||  
Db 517 VLVCAPSNIAVDOL 530

RESULT 13  
US-09-270-984A-2  
Sequence 2, Application US/09270984A  
Patent No. 6048965  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
TITLE OF INVENTION: NONSENSE-MEDIATED RNA DEGRADATION  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/270,984A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/724,354  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Hallie, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1118 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-270-984A-2

Query Match 60.5%; Score 34.5; DB 3; Length 1118;  
Best Local Similarity 64.3%; Pred. No. 1.4e+02;  
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 2 ILXCA-SNXAVDXL 14  
:|||||  
Db 517 VLVCAPSNIAVDOL 530

RESULT 14  
US-09-949-016-10116  
Sequence 10116, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: Venter, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CLO01307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10116  
LENGTH: 1140  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-10116

Query Match 60.5%; Score 34.5; DB 4; Length 1140;  
Best Local Similarity 64.3%; Pred. No. 1.4e+02;  
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 2 ILXCA-SNXAVDXL 14  
:|||||  
Db 539 VLVCAPSNIAVDOL 552

RESULT 15  
US-09-252-991A-16855  
Sequence 16855, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubinfeld et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 16855  
LENGTH: 259  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16855

Query Match 59.6%; Score 34; DB 4; Length 259;  
Best Local Similarity 50.0%; Pred. No. 33;  
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ILXCA-SNXAVDXL 14  
:|||||  
Db 168 RIAYCASKAGIDNL 181

Search completed: April 18, 2005, 08:18:32  
Job time: 18.4745 secs

This Page Blank (uspto)

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 08:06:16 ; Search time 49.6642 Seconds  
(without alignments)  
93.693 Million cell updates/sec

Title: US-10-652-334-2  
Perfect score: 57  
Sequence: 1 RILKASXAVDXL 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	89.5	14	US-10-652-334-2	Sequence 2, Appl1
2	40.5	71.1	380	US-10-652-334-29	Sequence 29, Appl1
3	40.5	71.1	611	US-10-369-493-10288	Sequence 10288, A
4	40.5	71.1	828	US-10-144-194A-96	Sequence 96, Appl1
5	40.5	71.1	1975	US-10-437-963-140079	Sequence 140079, A
6	40.5	71.1	2677	US-10-144-194A-22	Sequence 22, Appl1
7	38.5	65.8	712	US-10-369-493-22762	Sequence 22762, A
8	37.5	65.8	13	US-10-652-334-13	Sequence 13, Appl1
9	36.5	64.0	139	US-10-767-701-36594	Sequence 36594, A
10	36.5	64.0	472	US-10-652-334-26	Sequence 26, Appl1
11	36.5	64.0	626	US-10-437-963-166322	Sequence 166322, A
12	36.5	64.0	637	US-10-424-599-233501	Sequence 233501, A
13	36.5	64.0	639	US-10-425-114-37717	Sequence 37717, A

14	36.5	64.0	642	15	US-10-425-114-37557	Sequence 37557, A
15	36.5	64.0	2231	15	US-10-369-493-1830	Sequence 1830, Ap
16	35.5	62.3	262	15	US-10-424-599-252174	Sequence 252174, A
17	35.5	62.3	325	15	US-10-424-599-153343	Sequence 153343, A
18	35.5	62.3	992	15	US-10-369-493-12790	Sequence 12790, A
19	35.5	62.3	1944	15	US-10-369-493-2521	Sequence 2521, Ap
20	35	61.4	658	16	US-10-322-596-120	Sequence 144289, A
21	35	61.4	844	16	US-10-437-963-144289	Sequence 54346, A
22	35	61.4	880	15	US-10-425-114-54346	Sequence 38434, A
23	34.5	60.5	430	15	US-10-767-701-38434	Sequence 139, App
24	34.5	60.5	417	15	US-10-116-275-139	Sequence 8, Appl1
25	34.5	60.5	517	14	US-10-077-583-8	Sequence 2, Appl1
26	34.5	60.5	550	14	US-10-077-583-2	Sequence 3161, Ap
27	34.5	60.5	804	15	US-10-104-047-1161	Sequence 2087, Ap
28	34.5	60.5	925	15	US-10-369-493-2087	Sequence 1706, Ap
29	34.5	60.5	935	16	US-10-408-765A-1706	Sequence 6, Appl1
30	34.5	60.5	1118	16	US-10-474-553-6	Sequence 14, Appl1
31	34.5	60.5	1187	9	US-09-801-574-14	Sequence 4, Appl1
32	34.5	60.5	1208	14	US-10-077-583-4	Sequence 60, Appl1
33	34.5	60.5	1211	9	US-09-801-574-60	Sequence 111794, A
34	34.5	60.5	1323	16	US-10-437-963-111794	Sequence 6, Appl1
35	34.5	60.5	1349	14	US-10-077-583-6	Sequence 13091, A
36	34	59.6	128	16	US-10-437-963-193091	Sequence 234998, A
37	34	59.6	200	15	US-10-424-599-234998	Sequence 193088, A
38	34	59.6	516	16	US-10-437-963-193088	Sequence 115342, A
39	34	59.6	1969	16	US-10-437-963-115342	Sequence 219275, A
40	33	57.9	101	15	US-10-424-599-219275	Sequence 7104, Ap
41	33	57.9	243	15	US-10-369-493-7104	Sequence 4349, Ap
42	33	57.9	245	15	US-10-369-493-4349	Sequence 8151, Ap
43	33	57.9	330	14	US-10-156-761-8151	Sequence 69276, A
44	33	57.9	341	15	US-10-282-122A-69276	Sequence 264222, A
45	33	57.9	355	15	US-10-424-599-264222	

## ALIGNMENTS

RESULT 1  
US-10-652-334-2  
Sequence 2, Application US/10652334  
Publication No. US20040115787A1  
GENERAL INFORMATION:  
APPLICANT: Pelcz, Stuart  
APPLICANT: Czaplinski, Kevin  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/10/652,334  
CURRENT FILING DATE: 2003-08-28  
PRIOR APPLICATION NUMBER: US/09/359,268A  
PRIOR FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
PRIOR FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
FEATURE:  
OTHER INFORMATION: Xaa = any amino acid  
US-10-652-334-2

Query Match 89.5%; Score 51; DB 16; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.0027;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILKASXAVDXL 14  
DB 1 RILKASXAVDXL 14

```
RESULT 2
US-10-652-334-29
; Sequence 29, Application US/10652334
; Publication No. US20040115787A1
; GENERAL INFORMATION:
; APPLICANT: Pelcz, Stuart
; APPLICANT: Czaplinski, Kevin
; APPLICANT: Dimman, Jonathan D.
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
; FILE REFERENCE: 601-1-85N
; CURRENT APPLICATION NUMBER: US/10/652,334
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/359,268A
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,685
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-652-334-29

Query Match      71.1% Score 40.5; DB 16; Length 380;
Best Local Similarity 73.3%; Pred. No. 9.4;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY      1 RILXCA-SNXAVDXL 14
      ||| ||| ||| |||
Db      31 RILVCAPSNVAVDHL 45

RESULT 3
US-10-369-493-10288
; Sequence 10288, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10288
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Cytophaga hutchinsonii
US-10-369-493-10288

Query Match      71.1% Score 40.5; DB 15; Length 611;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY      1 RILXCA-SNXAVDXL 14
      ||| ||| ||| |||
Db      215 RVLVCASSNAVDVL 229

RESULT 4
US-10-144-194A-96
; Sequence 96, Application US/10144194A
; Publication No. US20030215809A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 96
; LENGTH: 828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-194A-96

Query Match      71.1% Score 40.5; DB 15; Length 828;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY      1 RILXCA-SNXAVDXL 14
      ||| ||| ||| |||
Db      153 RVLVCAPSNVAVDEL 167

RESULT 5
US-10-437-963-140079
; Sequence 140079, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 140079
; LENGTH: 1975
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_41310C.1.pep
US-10-437-963-140079

Query Match      71.1% Score 40.5; DB 16; Length 1975;
Best Local Similarity 66.7%; Pred. No. 50;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY      1 RILXCA-SNXAVDXL 14
      ||| ||| ||| |||
Db      1170 RVLVCAPSNVAVDEL 1184

RESULT 6
US-10-144-194A-22
; Sequence 22, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 2677
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4303)..(4305)
; OTHER INFORMATION: Xaa=Valine or Isoleucine
US-10-144-194A-22
```

```

Query Match          71.1%; Score 40.5; DB 15; Length 2677;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
```

```

Qy      1 RILXCA-SNXAVDXL 14
        ||| ||| ||| |||
Db      2002 RVLVCAPSNAVDL 2016
```

```

RESULT 7
US-10-369-493-22762
; Sequence 22762, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22762
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(712)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-22762
```

```

Query Match          67.5%; Score 38.5; DB 15; Length 712;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
```

```

Qy      1 RILXCA-SNXAVDXL 14
        ||| ||| ||| |||
Db      307 RILVCAPSNAVDL 321
```

```

RESULT 8
US-10-652-334-13
; Sequence 13, Application US/10652334
; Publication No. US20040115787A1
; GENERAL INFORMATION:
; APPLICANT: Peltz, Stuart
; APPLICANT: Czaplinski, Kevin
; APPLICANT: Dimman, Jonathan D.
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
; FILE REFERENCE: 601-1-85N
; CURRENT APPLICATION NUMBER: US/10/652,334
; PRIOR FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/359,268A
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,685
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 32
```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 13
; TYPE: PRT
; ORGANISM: saccharomyces cerevisiae
US-10-652-334-13
```

```

Query Match          65.8%; Score 37.5; DB 16; Length 13;
Best Local Similarity 76.9%; Pred. No. 1.2;
Matches 10; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
```

```

Qy      1 RILXCA-SNXAVD 12
        ||| ||| ||| |||
Db      1 RILVCAPSNIAVD 13
```

```

RESULT 9
US-10-767-701-36594
; Sequence 36594, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; PRIOR FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 36594
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: clone ID: SCORBI-28MAY03-C9756_1.pep
US-10-767-701-36594
```

```

Query Match          64.0%; Score 36.5; DB 16; Length 139;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;
```

```

Qy      1 RILXCA-SNXAVDXL 14
        ||| ||| ||| |||
Db      119 KILACASNTAVDNI 133
```

```

RESULT 10
US-10-652-334-26
; Sequence 26, Application US/10652334
; Publication No. US20040115787A1
; GENERAL INFORMATION:
; APPLICANT: Peltz, Stuart
; APPLICANT: Czaplinski, Kevin
; APPLICANT: Dimman, Jonathan D.
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
; FILE REFERENCE: 601-1-85N
; CURRENT APPLICATION NUMBER: US/10/652,334
; PRIOR FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/359,268A
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,685
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 472
; TYPE: PRT
; ORGANISM: saccharomyces cerevisiae
US-10-652-334-26
```

Query Match 64.0%; Score 36.5; DB 16; Length 472;  
Best Local Similarity 60.0%; Pred. No. 73;  
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

OY 1 RILKCA-ASNKAVDXL 14  
: || || || || || :  
DB 55 KILCAPSNNAVDI 69

RESULT 11  
US-10-437-963-166322  
; Sequence 166322, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 166322  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(626)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_65042C.1.pap  
US-10-437-963-166322

Query Match 64.0%; Score 36.5; DB 16; Length 626;  
Best Local Similarity 60.0%; Pred. No. 97;  
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

OY 1 RILKCA-ASNKAVDXL 14  
: || || || || || :  
DB 222 KILCAPSNNAVDI 236

RESULT 12  
US-10-424-599-233501  
; Sequence 233501, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 233501  
; LENGTH: 637  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(637)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_52878C.1.pap  
US-10-424-599-233501

Query Match 64.0%; Score 36.5; DB 15; Length 637;  
Best Local Similarity 60.0%; Pred. No. 99;  
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

OY 1 RILKCA-ASNKAVDXL 14  
: || || || || || :  
DB 234 KILCAPSNNAVDI 248

RESULT 13  
US-10-425-114-37717  
; Sequence 37717, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53131)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 37717  
; LENGTH: 639  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(639)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3170-056-H12\_F11.pap  
US-10-425-114-37717

Query Match 64.0%; Score 36.5; DB 15; Length 639;  
Best Local Similarity 60.0%; Pred. No. 99;  
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

OY 1 RILKCA-ASNKAVDXL 14  
: || || || || || :  
DB 234 KILCAPSNNAVDI 248

RESULT 14  
US-10-425-114-37557  
; Sequence 37557, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53131)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 37557  
; LENGTH: 642  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana columbia  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(642)  
; OTHER INFORMATION: Clone ID: LIB3175-019-D10\_F11.pap  
US-10-425-114-37557

Query Match 64.0%; Score 36.5; DB 15; Length 642;  
Best Local Similarity 60.0%; Pred. No. 99;

Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;  
 Qy 1 RILXC-ASXAVDXL 14  
 :|||  
 Db 238 KILCAPSNAVDNI 252

RESULT 15  
 US-10-369-493-1830  
 ; Sequence 1830, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; PRIOR FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 1830  
 ; LENGTH: 2231  
 ; TYPE: PRT  
 ; ORGANISM: Saccharomyces cerevisiae  
 US-10-369-493-1830

Query Match 64.0%; Score 36.5; DB 15; Length 2231;  
 Best Local Similarity 60.0%; Pred. No. 3.5e+02;  
 Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 1 RILXC-ASXAVDXL 14  
 :|||  
 Db 1405 KILCAPSNAVDNI 1419

Search completed: April 18, 2005, 09:04:05  
 Job time : 49.6642 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 18, 2005, 07:32:05 ; Search time 13.4891 Seconds  
(without alignments)  
99.861 Million cell updates/sec

Title: US-10-652-334-2

Perfect score: 57

Sequence: 1 RILXCASNVAVDXL 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40.5	71.1	971	2	S23408
2	40.5	71.1	2142	2	D86303
3	38.5	67.5	660	2	T41580
4	38	66.7	433	2	B87153
5	38	66.7	433	2	B70896
6	38	66.7	1671	2	T31628
7	37	64.9	1173	2	T30608
8	36.5	64.0	635	2	T02699
9	36.5	64.0	989	2	T48845
10	36.5	64.0	1825	2	T52521
11	36.5	64.0	2231	2	S53416
12	35.5	62.3	555	2	E71420
13	35.5	62.3	993	2	A47500
14	35.5	62.3	993	2	S35633
15	35.5	62.3	1944	2	T40065
16	34.5	60.5	678	2	T42668
17	34.5	60.5	935	2	S62476
18	34.5	60.5	1048	2	C63189
19	34.5	60.5	1090	2	T00533
20	34.5	60.5	1311	2	T08986
21	34.5	60.5	1687	2	T39072
22	34.5	60.5	1687	2	B82544
23	34	59.6	234	2	C63217
24	34	59.6	234	2	AH2454
25	33	57.9	274	2	B95500
26	33	57.9	280	2	D97185
27	33	57.9	488	2	S40706
28	33	57.9	497	2	S37580
29	33	57.9	497	2	S37581

30	33	57.9	537	1	JU0150	cellulose 1,4-beta
31	33	57.9	635	2	S73017	polyketide synthase
32	33	57.9	873	2	B87049	conserved hypothet
33	33	57.9	875	2	P70755	hypothetical prote
34	33	57.9	1446	2	S73013	polyketide synthase
35	33	57.9	1540	2	B87203	polyketide synthase
36	33	57.9	1871	2	A87204	polyketide synthase
37	32.5	57.0	683	2	S34700	probable purine nu
38	32	56.1	72	2	S75319	hypothetical prote
39	32	56.1	124	2	S01369	hypothetical prote
40	32	56.1	133	1	HANERN	hemoglobin alpha c
41	32	56.1	346	2	AB0942	hypothetical prote
42	32	56.1	381	2	S45766	hypothetical prote
43	32	56.1	457	2	B96838	hypothetical prote
44	32	56.1	607	2	P82339	conserved hypothet
45	32	56.1	659	2	A57532	Na+-dependent nucl

## ALIGNMENTS

## RESULT 1

S23408 prematurely terminated mRNA decay factor NAM7 - yeast (Saccharomyces cerevisiae)

N/Alternate names: protein YM9583.05c; protein YMR080c; UPF1 protein

C/Species: Saccharomyces cerevisiae

C/Date: 04-Dec-1992 #sequence revision 04-Dec-1992 #text\_change 09-Jul-2004

C/Accession: S23408; A44388; S54455

R/Altamura, N.; Grudinsky, O.; Dujardin, G.; Slonimski, P.P.

J. Mol. Biol. 224, 575-587, 1992

A/Title: NAM7 nuclear gene encodes a novel member of a family of helicases with a Zn-11g

A/Reference number: S23408; MUID:92235815; PMID:1314899

A/Accession: S23408

A/Molecule type: DNA

A/Residues: 1-971 <ALT>

A/Cross-references: UNIPROT:P30771; EMBL:X62394; NID:94022; PIDN:CAA44266.1; PID:94023

R/Gentiles, S.; Bowman, S.

Submitted to the EMBL Data Library, May 1995

A/Reference number: S54451

A/Accession: S54455

A/Molecule type: DNA

A/Residues: 1-971 <GEN>

A/Cross-references: EMBL:Z49259; NID:9807956; PID:9807962; MIPS:YMR080C

A/Experimental source: Ectrain AB972

C/Genetics:

A/Gene: SGD:NAM7; UPF1

A/Cross-references: SGD:S0004685; MIPS:YMR080C

A/Map position: 13R

C/Keywords: GTP binding; microtubule-binding; nucleotide binding; nucleus; P-loop

F/430-437/Region: nucleotide-binding motif A (P-loop)

F/548-548/Region: GTP-binding NKXD motif

Query Match 71.1% Score 40.5; DB 2; Length 971;

Best Local Similarity 73.3% Pred. No. 3.3;

Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 RILXCA-SNNAVDXL 14

Db 454 RILVCAPSNVAVDHL 468

## RESULT 2

D86303 P17F16.1 protein - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

A:Accession: B87153  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-433 <STO>  
A:Cross-references: UNIPROT:Q9X792; GB:AL450380; NID:g13093606; PIDN:CA030906.1; GSPDB:GT  
C:Genetics:  
A:Gene: phoH2

Query Match 66.7%; Score 38; DB 2; Length 433;  
Best Local Similarity 58.3%; Pred. No. 5.2;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RILXCASMXAVD 12  
||| ||| : :  
Db 106 RILSCAATLAAR 117

RESULT 5  
E70896  
probable phoH2 protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
A:Accession: E70896  
R:CoLe, S.T.; Brosh, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
C:Conom, R.; Davies, R.; Devlin, K.; Felwell, T.; Genies, S.; Hamlin, N.; Holroyd, S.;  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: E70896  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-433 <COL>  
A:Cross-references: UNIPROT:O53443; GB:AL021897; GB:AL123456; NID:g3256022; PIDN:CAA17211  
C:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: phoH2

Query Match 66.7%; Score 38; DB 2; Length 433;  
Best Local Similarity 58.3%; Pred. No. 5.2;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RILXCASMXAVD 12  
||| ||| : :  
Db 106 RILSCAATLAAR 117

RESULT 6  
E71628  
sensory transduction histidine kinase dokA - slime mold (Dictyostelium discoideum)  
C:Species: Dictyostelium discoideum  
C:Date: 27-Nov-1997 #sequence\_revision 12-Dec-1997 #text\_change 09-Jul-2004  
A:Accession: E71628; S78068  
R:Schuster, S.C.; Noegel, A.A.; Oehme, F.; Gerlsch, G.; Simon, M.I.  
EMBO J. 15, 3880-3889, 1996  
A:Title: The hybrid histidine kinase DokA is part of the osmotic response system of Dicty  
A:Reference number: E71628; MUID:96324396; PMID:8670893  
A:Accession: E71628  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-1670 <SCH>  
A:Cross-references: UNIPROT:Q23901; EMBL:X96869  
A:Experimental source: strain AX2; substrain 214  
R:Schuster, S.C.; Noegel, A.A.; Oehme, F.; Gerlsch, G.; Simon, M.I.  
submitted to: The EMBL Data Library, March 1996  
A:Description: The hybrid histidine kinase DokA is part of the osmotic response system of  
A:Reference number: S78068  
A:Residues: 1-149; 'E', 151-219, 'TRVLKLIQSTNNMIVY', 238-1671 <SCW>  
A:Cross-references: EMBL:X96869; NID:g1237201; PIDN:CA05612.1; PID:g1237202  
C:Genetics:



A:Gene: dokA

C:Function:

A:Description: modulates cell response to changes in osmolarity; involved in spore formation  
 C:Keywords: phosphoprotein; signal transduction  
 F:1550-1629/Domain: response regulator homology <RRH2>  
 F:1568/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match

Best Local Similarity 66.7%; Score 38; DB 2; Length 1671;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ILKCSNXXAVDXL 14

Db 298 ILKCSNXXAVDXL 310

RESULT 7  
 T30608  
 proteophosphoglycan homolog 6L - Molluscum contagiosum virus 1

N:Alternate names: MC006L

C:Species: Molluscum contagiosum virus 1

C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004

C:Accession: T30608

R:Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.

A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re

A:Reference number: Z20876; MUID:96325459; PMID:8670425

A:Accession: T30608

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Cross-references: UNIPROT:Q98177; EMBL:U60315; PIDN:AAC55134.1

C:Genetics:

A:Note: MC006L

Query Match

Best Local Similarity 64.9%; Score 37; DB 2; Length 1173;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ILKCSNXXAVDXL 12

Db 1051 ILKCSNXXAVDXL 1061

RESULT 8  
 T02699  
 probable helicase At2g03270 [imported] - Arabidopsis thaliana

N:Alternate names: hypothetical protein T18E12.6

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004

C:Accession: T02699; D84446

R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Rot

A:Description: The Arabidopsis thaliana chromosome II BAC T18E12 genomic sequence.

A:Reference number: Z14702

A:Accession: T02699

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-635 &lt;ROU&gt;

A:Cross-references: UNIPROT:O81047; EMBL:AC005313; NID:G3548797; PID:G3548803

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Wolf, K.S.; Cronin, L.A.; Shen, M.; Yankem, S.E.; Umayam, L.; Tailon, L.

euser, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-769, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: D84446

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-635 &lt;STO&gt;

A:Cross-references: GB:AE002093; NID:G4335770; PIDN:AA017447.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g03270; T18E12.6

A:Map position: 2

C:Superfamily: probable DNA helicase MJ0104

Query Match

Best Local Similarity 64.0%; Score 36.5; DB 2; Length 635;  
 Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 ILKCA-SNXXAVDXL 14

Db 231 ILKCA-SNXXAVDXL 245

RESULT 9  
 T48845  
 insulin II gene enhancer-binding protein rip1 [validated] - golden hamster

C:Species: Mesocricetus auratus (golden hamster)

C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004

C:Accession: T48845

R:Shieh, S.Y.; Stellrecht, C.M.; Teal, M.J.

J. Biol. Chem. 270, 21503-21508, 1995

A:Title: Molecular characterization of the rat insulin enhancer-binding complex 3b2. Clo

A:Reference number: Z24545; MUID:95394901; PMID:7665561

A:Accession: T48845

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: RNA

A:Residues: 1-989 &lt;SH&gt;

A:Cross-references: UNIPROT:Q60560; EMBL:L15625; NID:G2790918; PIDN:AA00104.1; PID:G2909

A:Experimental source: tissue type insulinoma

C:Genetics:

A:Note: RIP1

C:Complex: rip1 is a component of the RIP3b-binding complex 3b2, which is composed of a

A:Description: the 3b2 complex is one of two complexes, binding to the RIP3b element of

ent RIP3a [validated, MUID:95394901]

Query Match

Best Local Similarity 64.0%; Score 36.5; DB 2; Length 989;  
 Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 1 ILKCA-SNXXAVDXL 14

Db 236 ILKCA-SNXXAVDXL 250

RESULT 10  
 T52521  
 related to SEN1 protein [imported] - Neurospora crassa

N:Alternate names: protein B2J23.170

C:Species: Neurospora crassa

C:Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004

C:Accession: T52521

R:Schulze, U.; Aism, V.; Hohelsel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, September 2000

A:Reference number: Z26053

A:Accession: T52521

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1825 &lt;SCH&gt;

A:Cross-references: UNIPROT:Q9HFI5; EMBL:AL442164; GSPDB:GN00116; NCSP:B2J23.170

A:Experimental source: BAC clone B2J23; strain OR74A

C:Genetics:

A:Gene: NCSP:B2J23.170

A:Map position: 6

Query Match

Best Local Similarity 64.0%; Score 36.5; DB 2; Length 1825;  
 Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 ILKCA-SNXXAVDXL 14

Db 1218 KILVCA-SNXXAVDXL 1232

## RESULT 11

SEN1 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein L9576.1; protein YLR430w

C:Species: Saccharomyces cerevisiae

C>Date: 05-May-1995 #sequence revision 01-Sep-1995 #text\_change 09-Jul-2004

C:Accession: S53416; A44387; S41985

R:Reviewed: A.

Submitted to the EMBL Data Library, February 1995

A:Description: The sequence of *S. cerevisiae* comid 9576.

A:Reference number: S53409

A:Accession: S53416

A:Molecule type: DNA

A:Residues: 1-2231 <PRV>

A:Cross-references: UNIPROT:Q00416; EMBL:U20939; NID:g664871; PID:g664872; MIPS:YLR430w

A:Experimental source: strain S28C (AB972)

R:demarini, D.J.; Winey, M.; Ursic, D.; Webb, F.; Culbertson, M.R.

Mol. Cell. Biol. 12, 2154-2164, 1992

A:Title: SEN1, a positive effector of tRNA-splicing endonuclease in *Saccharomyces cerevisiae*

A:Reference number: A44387; MUID:92236590; PMID:1565945

A:Accession: A44387

A:Molecule type: DNA

A:Residues: 'MMS', 130, 'PCERVO', 131-2231 <DEM>

A:Cross-references: GB:M74589; NID:g172573; PID:AB63976.1; PID:g172574

C:Genetics:

A:Gene: SCD,SEN1

A:Cross-references: SCD:S0004422; MIPS:YLR430w

A:Map position: 12R

C:Function:

A:Description: may be component of nuclear splicing complex

C:Keywords: nucleotide binding; nucleus; P-loop

F:1357-1364/Region: nucleotide-binding motif A (P-loop)

Query Match

Best Local Similarity 64.0%; Score 36.5; DB 2; Length 2231;

Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

1 RILKCA-SNXAVDXL 14

1405 KILICAPSNAAVDI 1419

## RESULT 12

hypothetical protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: Columbia

C>Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 09-Jul-2004

C:Accession: E71420

R:Beran, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk

P.; Wedler, H.; Wedler, E.; Wandutt, R.; Weitzenecker, T.; Pohl, T.M.; Terryn, N.; Giel

avannah, T.; Hempel, S.; Kotter, P.; Brittan, K.D.; Rieger, M.; Schaefer, M.; Funk, B.

Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech

erhoth, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Reichman, S.; Ans

C.; Chavarzie, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thal*

A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: E71420

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-555 <BV>

A:Cross-references: UNIPROT:O23408; GB:Z97339; NID:g2244901; PID:e326936; PID:g2244913

C:Genetics:

A:Map position: 4COP9-4G3845

Query Match

Best Local Similarity 62.3%; Score 35.5; DB 2; Length 555;

Matches 8; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

1 RILKCA-SNXAVDXL 14

1 RILKCA-SNXAVDXL 14

Db 109 RVLICAPSNALDEI 123

## RESULT 13

19 mu chain switch region binding protein 2 - human

N:Alternate names: DNA-binding protein 8-mu-bp-2; glial factor 1; transcription factor G1

C:Species: Homo sapiens (man)

C>Date: 02-Jun-1995 #sequence\_revision 02-Jun-1995 #text\_change 09-Jul-2004

C:Accession: A47500; A40804

R:Fukita, Y.; Mizuta, T.R.; Shirozu, M.; Ozawa, K.; Shimizu, A.; Honjo, T.

J. Biol. Chem. 268, 17463-17470, 1993

A:Title: The human Smdbp-2, a DNA-binding protein specific to the single-stranded guanine

A:Reference number: A47500; MUID:93352537; PMID:8349627

A:Accession: A47500

A:Molecule type: mRNA

A:Residues: 1-993 <FK>

A:Cross-references: UNIPROT:P38935; GB:L14754

R:Kerr, D.; Khalil, K.

J. Biol. Chem. 266, 15876-15881, 1991

A:Title: A recombinant cDNA derived from human brain encodes a DNA binding protein that

A:Reference number: A40804; MUID:91340730; PMID:1714899

A:Accession: A40804

A:Molecule type: mRNA

A:Residues: 'GGRV', 495-862, 'K', 864-865, 'T' <KER>

A:Cross-references: GB:M64979; NID:g183249; PID:AA58611.1; PID:g183250

A:Experimental source: brain stem

A>Note: the authors translated the codon CAG for residue 507 as Glu

C:Genetics:

A:Gene: GDB:IGHMBP2

A:Cross-references: GDB:434023; OMIM:600502

A:Map position: 11q13.2-11q13.4

C:Keywords: immunoglobulin; single-stranded DNA binding

Query Match

Best Local Similarity 62.3%; Score 35.5; DB 2; Length 993;

Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

1 RILKCA-SNXAVDXL 14

237 KVLICAPSNIAVDNL 251

## RESULT 14

DNA-binding protein - mouse

C:Species: Mus musculus (house mouse)

C>Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004

C:Accession: S35633

R:Mizuta, T.R.; Fukita, Y.; Miyoshi, T.; Shimizu, A.; Honjo, T.

Nucleic Acids Res. 21, 1761-1766, 1993

A:Title: Isolation of cDNA encoding a binding protein specific to 5'-phosphorylated singl

A:Reference number: S35633; MUID:93261806; PMID:8493094

A:Accession: S35633

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-993 <MT>

A:Cross-references: UNIPROT:P40694; GB:L10075; NID:g293805; PID:AAA40143.1; PID:g293806

C:Keywords: DNA binding

Query Match

Best Local Similarity 62.3%; Score 35.5; DB 2; Length 993;

Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

1 RILKCA-SNXAVDXL 14

236 KVLICAPSNIAVDNL 250

## RESULT 15

crRNA-splicing endonuclease positive effector - fission yeast (*Schizosaccharomyces pombe*)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: T40065  
R/Lyne, M.; Rafandream, M.A.; Barrell, B.G.; Lelaure, V.; Galibert, F.  
submitted to the EMBL Data Library, December 1998  
A/Reference number: Z21903  
A/Accession: T40065  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1944 <LVN>  
A/Cross-references: UNIPROT:O94387; EMBL:AI034463; P1DN:CAA22438.1; GSPDB:GN00067; SPDB:  
A/Experimental source: strain 972h-; coemid c29A10  
C/Genetics:  
A/Gene: SPDB:SPBC29A10.10c  
A/Map position: 2

Query Match 62.3%; Score 35.5; DB 2; Length 1944;  
Best Local Similarity 53.3%; Pred. No. 63;  
Matches 8; Conservative 3; Mismatches 3; Indels 1; Gaps 1;  
Qy 1 RILXCA-SRXAVDXL 14  
: ||| ||| :  
Db 1333 KILICPSPNAIDBI 1347

Search completed: April 18, 2005, 08:05:59  
Job time : 15.4891 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 18, 2005, 07:33:55 ; Search time 58.4526 Seconds  
(without alignments)  
122.648 Million cell updates/sec

Title: US-10-652-334-2  
Perfect score: 57  
Sequence: 1 RILKASXAVDXL 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues  
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41.5	72.8	763	Q7UWPI	Q7UWPI rhodopirell
2	40.5	71.1	778	Q6Z081	Q6Z081 mus musculu
3	40.5	71.1	821	Q6PEDE	Q6PEDE mus musculu
4	40.5	71.1	902	Q8OV90	Q8OV90 mus musculu
5	40.5	71.1	917	Q6AZD0	Q6AZD0 homo sapien
6	40.5	71.1	964	Q6FVZ1	Q6FVZ1 candida gla
7	40.5	71.1	971	NAM7_YEAST	P30771 saccharomyc
8	40.5	71.1	2142	Q6FWR3	Q6FWR3 arabidopsis
9	40.5	71.1	2198	Q6SDWS	Q6SDWS homo sapien
10	40.5	71.1	2646	Q6IMG6	Q6IMG6 mus musculu
11	40.5	71.1	2677	SERTX_HUMAN	Q73333 homo sapien
12	39.5	69.3	969	Q6CWM6	Q6CWM6 kluyveromyc
13	39.5	69.4	1000	Q75DS7	Q75DS7 ashyba go88
14	39	68.4	191	Q8B860	Q8B860 pseudomonas
15	38.5	67.5	611	Q7PP43	Q7PP43 anophelies g
16	38.5	67.5	660	Q94247	Q94247 schizosach
17	38.5	67.5	695	Q6FKP9	Q6FKP9 candida gla
18	38.5	67.5	967	Q6BAS0	Q6BAS0 dictyostell
19	38.5	67.5	1968	Q6BWT0	Q6BWT0 debaromyce
20	38	66.7	433	Q53443	Q53443 mycobacteri
21	38	66.7	433	Q73WQ3	Q73WQ3 mycobacteri
22	38	66.7	433	Q7U0N9	Q7U0N9 mycobacteri
23	38	66.7	433	Q9X792	Q9X792 mycobacteri
24	38	66.7	1670	Q23901	Q23901 dictyostell
25	38	66.7	1671	Q6B9S5	Q6B9S5 dictyostell
26	37.5	65.8	734	Q7OYL6	Q7OYL6 giardia lam
27	37.5	65.8	964	Q6C803	Q6C803 yarrowia li
28	37.5	65.8	985	Q6BPM3	Q6BPM3 debaromyce
29	37.5	65.8	1997	Q6CWA6	Q6CWA6 kluyveromyc
30	37.5	65.8	2027	Q756Z8	Q756Z8 ashyba go88
31	37	64.9	227	O11286	O11286 molluscum c

32	37	64.9	1173	2	Q98177	Q98177 molluscum c
33	37	64.9	1175	2	Q85285	Q85285 molluscum c
34	36.5	64.0	635	2	Q9FNX9	Q9FNX9 arabidopsis
35	36.5	64.0	639	2	O81047	O81047 arabidopsis
36	36.5	64.0	651	2	Q6F2U8	Q6F2U8 oryza sativ
37	36.5	64.0	989	1	SMB2_MESAU	Q60560 mesocricetu
38	36.5	64.0	1120	2	Q7PWZ4	Q7PWZ4 anophelies g
39	36.5	64.0	1417	2	Q9VSE1	Q9VSE1 drosophila
40	36.5	64.0	1979	2	Q6FK55	Q6FK55 candida gla
41	36.5	64.0	2126	2	Q9HPT5	Q9HPT5 neurospora
42	36.5	64.0	2231	1	SENI_YEAST	Q00416 saccharomyc
43	36	63.2	1065	1	PRR3_ARATH	Q91V94 arabidopsis
44	36	63.2	1065	2	Q72V06	Q72V06 leptospira
45	36	63.2	1065	2	Q8EZX0	Q8EZX0 leptospira

## ALIGNMENTS

RESULT 1						
ID	Q7UWPI	PRELIMINARY;	PRT;	763 AA.		
AC	Q7UWPI					
DT	01-OCT-2003 (TREMBLrel. 25, Created)					
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)					
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)					
DE	DNA-binding protein SMUBP-2.					
GN	OrderedLocustNames=RB1895;					
OS	Rhodopirellula Ballica.					
OC	Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;					
OC	Planctomycetaceae; Pirellula.					
OX	NCBI_TaxID=117;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRUIN=1;					
RX	MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;					
RA	Glockner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,					
RA	Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,					
RA	Schneider H., Mann R., Reinhardt R.;					
RT	"Complete genome sequence of the marine planctomycete Pirellula sp.					
RT	strain 1."					
RL	Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).					
DR	EMBL; BX294136; CAD72321.1; -					
DR	GO; GO:0005524; F:ATP binding; IEA.					
DR	GO; GO:0003677; F:DNA binding; IEA.					
DR	GO; GO:0000166; F:nucleotide binding; IEA.					
DR	InterPro; IPR003593; AAA_ATPase.					
DR	InterPro; IPR011545; DEAD/DEAF_N.					
DR	SMART; SM00382; AAA; 1.					
DR	SMART; SM00487; DEXDC; 1.					
KW	ATP-binding; Complete proteome; DNA-binding.					
SQ	SEQUENCE 763 AA; 84494 MW; ASDA57C25BDBCC CRC64;					
Query Match						
		72.8%;	Score 41.5;	DB 2;	Length 763;	
		Best Local Similarity	66.7%;	Pred. No. 6.5;		
		Matches 10;	Conservative 1;	Mismatches 3;	Indels 1;	Gaps 1;
Qy	1 RILK-ASXAVDXL 14					
Db	356 RILKASXAVDXL 370					
RESULT 2						
ID	Q6Z081	PRELIMINARY;	PRT;	778 AA.		
AC	Q6Z081					
DT	05-JUL-2004 (TREMBLrel. 27, Created)					
DT	05-JUL-2004 (TREMBLrel. 27, Last sequence update)					
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)					
DE	MK1A0625 protein (Fragment).					
GN	Name=MK1A0625;					
OS	Mus musculus (Mouse).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryonic tail;  
 RX PubMed=14621295;  
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,  
 RA Suga Y., Nagase T., Ohara O., Koga H.;  
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:  
 RT 111. the complete nucleotide sequences of 500 mouse KIAA-homologous  
 RT cDNAs identified by screening of terminal sequences of cDNA clones  
 RT randomly sampled from size-fractionated libraries.";  
 RL DNA Res. 10:167-180(2003).  
 DR EMBL; AK129177; BAC97987.1; -  
 FT NOW TRR 1  
 SQ SEQUENCE 778 AA; 87017 MW; 51D5E3DAAB3AF4EA CRC64;  
 Query Match 71.1%; Score 40.5; DB 2; Length 778;  
 Best Local Similarity 66.7%; Pred. No. 11;  
 Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
 Oy 1 RILXCA-SNXAVDXL 14  
 Db 64 RVLVCAPSNAVDEL 78  
 RESULT 3  
 ID Q6PED8 PRELIMINARY; PRT; 821 AA.  
 AC Q6PED8;  
 DT 05-JUL-2004 (TRENBLREL. 27, Created)  
 DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)  
 DT 25-OCT-2004 (TRENBLREL. 28, Last annotation update)  
 DE Expressed sequence AW060766.  
 GN Name=AW060766;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129; TISSUE=Mammary tumor. Brca1-/fl;  
 RA Strauberg R.;  
 RL Submitted (FEB-2003) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; BC046382; AA46382.1; -  
 DR MGD; MGI:2139292; AW060766.  
 FT NOW TRR 1  
 SQ SEQUENCE 902 AA; 101988 MW; E4356FB7063D94F CRC64;  
 Query Match 71.1%; Score 40.5; DB 2; Length 902;  
 Best Local Similarity 66.7%; Pred. No. 12;  
 Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
 Oy 1 RILXCA-SNXAVDXL 14  
 Db 234 RVLVCAPSNAVDEL 248  
 RESULT 5  
 ID Q6AZD7 PRELIMINARY; PRT; 917 AA.  
 AC Q6AZD7;  
 Submitted (AUG-2004) to the EMBL/Genbank/DBJ databases.

DR EMBL; BC058109; AA58109.1; -  
 DR EMBL; BC079604; AA479604.1; -  
 SQ SEQUENCE 821 AA; 92120 MW; CB28DFCA91698287 CRC64;  
 Query Match 71.1%; Score 40.5; DB 2; Length 821;  
 Best Local Similarity 66.7%; Pred. No. 11;  
 Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
 Oy 1 RILXCA-SNXAVDXL 14  
 Db 153 RVLVCAPSNAVDEL 167  
 RESULT 4  
 ID Q8OV90 PRELIMINARY; PRT; 902 AA.  
 AC Q8OV90;  
 DT 01-JUN-2003 (TRENBLREL. 24, Created)  
 DT 01-JUN-2003 (TRENBLREL. 24, Last sequence update)  
 DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)  
 DE AM060766 protein (Fragment).  
 GN Name=AM060766;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129; TISSUE=Mammary tumor. Brca1-/fl;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129; TISSUE=Mammary tumor. Brca1-/fl;  
 RA Strauberg R.;  
 RL Submitted (FEB-2003) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; BC046382; AA46382.1; -  
 DR MGD; MGI:2139292; AW060766.  
 FT NOW TRR 1  
 SQ SEQUENCE 902 AA; 101988 MW; E4356FB7063D94F CRC64;  
 Query Match 71.1%; Score 40.5; DB 2; Length 902;  
 Best Local Similarity 66.7%; Pred. No. 12;  
 Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
 Oy 1 RILXCA-SNXAVDXL 14  
 Db 234 RVLVCAPSNAVDEL 248  
 RESULT 5  
 ID Q6AZD7 PRELIMINARY; PRT; 917 AA.  
 AC Q6AZD7;  
 Submitted (AUG-2004) to the EMBL/Genbank/DBJ databases.

```

DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DB Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Alechul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepieten M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
RA Fahey J., Hellon E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S.,
RA Krzywinski M.I., Skalska U., Small D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RA Strausberg R.;
RL Submitted (JUL-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC078166; AAH78166.1; -.
KM Hypothetical protein.
FT NON_TER
SQ SEQUENCE 917 AA; 104224 MW; AB3C550CFPA29ABA CRC64;

Query Match 71.1%; Score 40.5; DB 2; Length 917;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Cy 1 RILKCA-SNKAVDXL 14
Db 242 RILVCAPSNIAVDL 256

RESULT 6
O6FVZ1 PRELIMINARY; PRT; 964 AA.
AC O6FVZ1;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DB Candida glabrata strain CBS138 chromosome D complete sequence.
GN ORFNames=CAGL004312g;
OS Candida glabrata CBS138.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mtosporic Saccharomycetales; Candida.
OX NCBI_TaxID=284593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durans P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marcq C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anichoud V., Babor A., Barbe V.,
RA Barry S., Blanchin S., Beckerich J.M., Beyne R., Bleykasten C.,
RA Bolserne A., Boyer J., Catolico L., Confantolero F., de Darvar A.,

```

```

RA Despons L., Fabre R., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantreay F., Hemequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemire M., Lesur I., Ma L., Muller R.,
RA Nicoud J.M., Nikolaki M., Ozas S., Ozler-Kalogeropoulos O.,
RA Pellens S., Potier S., Richard G.P., Straub M.U., Suleau A.,
RA Swennens D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts."
RL Nature 430:35-44(2004).
DR EMBL: CR380950; CAG58514.1; -.
DR InterPro: IPR01545; DEAD/DEAH_N.
DR InterPro: IPR006935; Resili.
DR Pfam: PF04851; Resili; 1.
DR SMART: SM00487; DEXDC; 1.
SQ SEQUENCE 964 AA; 106274 MW; 0896F59C957596B CRC64;

Query Match 71.1%; Score 40.5; DB 2; Length 964;
Best Local Similarity 73.3%; Pred. No. 13;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Cy 1 RILKCA-SNKAVDXL 14
Db 453 RILVCAPSNIAVDL 467

RESULT 7
NM7_YEAST STANDARD; PRT; 971 AA.
AC P30771;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DB NM7 protein (Nuclear accommodation of mitochondria 7 protein)
DB Nonense-mediated mRNA decay protein 1) (Up-frameshift suppressor 1).
GN Name=NM7; Synonyms=IFS2, MOF4, UPF1; OrderedLocNames=YMR080C;
GN ORFNames=YMS582.05C;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetales; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R23/50;
RX MEDLINE=92235815; PubMed=1314899;
RA Altamura N., Groudinsky O., Dujardin G., Slonimski P.P.;
RT "NM7 nuclear gene encodes a novel member of a family of helicases
RT with a Zn-ligand motif and is involved in mitochondrial functions in
RT Saccharomyces cerevisiae."
RL J. Mol. Biol. 224:575-587(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII."
RL Nature 387:90-93(1997).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=97051830; PubMed=8896465;
RA Cui Y., Dimman J.D., Peltz S.W.;

```

```

RT      "MofA-1 is an allele of the UPF1/IFS2 gene which affects both mRNA
RT      turnover and -1 ribosomal frameshifting efficiency.";
RT      EMBL J. 15:5726-5736(1996).
CC      -1- FUNCTION: Probable helicase involved in mitochondrial functions.
CC      Required for rapid turnover of mRNAs containing a premature
CC      translational termination codon.
CC      -1- SUBCELLULAR LOCATION: Present predominantly in the cytoplasm, but
CC      is also found in small quantities in the nucleus.
CC      -1- SIMILARITY: Belongs to the DNM2/NAM7 helicase family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X62394; CAA44266.1; -.
DR      EMBL; M76659; AAA5197.1; -.
DR      EMBL; Z49259; CAA89226.1; -.
DR      PIR; S23408; S23408.
DR      GenBank; 142747; -.
DR      SGD; S000004685; NAM7.
DR      GO; GO:0005737; C:cytoplasm; IDA.
DR      InterPro; IPR003593; AAA ATPase.
DR      InterPro; IPR001410; DEAD.
DR      SMART; SM00382; AAA; 1.
DR      SMART; SM00487; DEXDC; 1.
KM      ATP-binding; Helicase; Hydrolyase; Mitochondrion;
KM      Nonsense-mediated mRNA decay; Nuclear protein; Zinc-finger.
FT      ZN FING 70 98 C2H2-type (atypical) (Potential).
FT      ZN FING 122 152 C4-type (Potential).
FT      NP BIND 430 437 ATP (By similarity).
SQ      SEQUENCE 971 AA; 109430 MW; 9161AFB0BE6747FE CRC64;

Query Match 71.1%; Score 40.5; DB 1; Length 971;
Best Local Similarity 73.3%; Pred. No. 13;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY      1 RILXCA-SNXAVDXL 14
DB      454 RILWCAPSNVAVDHL 468

RESULT 8
Q9FWR3 PRELIMINARY; PRT; 2142 AA.
AC Q9FWR3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE F17F16.1 protein.
GN Name=F17F16.1;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federapfel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Alalaf H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA Buehler E., Chao Q., Chin C., Chlou J., Choi E., Gonzalez A.,
RA Huang B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA Lenz C., Liu A., Liu S., Mukharbeky N., Pham P., Sakano H., Shim P.,
RA Tottum M., Vaynsberg M., Yu G., Becker J., Theologis A., Davis R.W.;
RA Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC026237; AAC09081.1; -.
DR PIR; D86303; D86303.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR SMART; SM00487; DEXDC; 1.
SQ SEQUENCE 2142 AA; 240657 MW; B6F6AA017A42CA2 CRC64;

```

```

Query Match 71.1%; Score 40.5; DB 2; Length 2142;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY      1 RILXCA-SNXAVDXL 14
DB      1306 RVLICQSNNAVDL 1320

RESULT 9
Q68DW5 PRELIMINARY; PRT; 2198 AA.
ID Q68DW5
AC Q68DW5;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp781B151 (Fragment).
GN Name=DKFZp781B151;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RG The German cDNA Consortium;
RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Oeanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR749249; CAH18105.1; -.
KM Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 2198 AA; 247685 MW; D9675C85AAE5A089 CRC64;

Query Match 71.1%; Score 40.5; DB 2; Length 2198;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY      1 RILXCA-SNXAVDXL 14
DB      1556 RVLWCAPSNNAVDL 1570

RESULT 10
Q6IMG6 PRELIMINARY; PRT; 2646 AA.
ID Q6IMG6
AC Q6IMG6;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Senataxin.
GN Name=Setx;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX PubMed=14770181; DOI=10.1038/ng1303;
RA Moreira M.C., Klur S., Watanabe M., Nemeth A.H., Le Ber I.,
RA Moritz J.C., Tranchant C., Aubourg P., Tazir M., Schols L.,
RA Pardo L., Schulz J.B., Pouget J., Calvas P., Shizuka-Ikeda M.,
RA Shoji M., Tanaka M., Izat L., Shaw C.B., M'Zahem A., Dunne E.,
RA Bonnot P., Benhassine T., Bouslam N., Stevanin G., Brice A.,
RA Gammaraes J., Mendonca P., Barbot C., Coutinho P., Sequeiros J.,
RA Durr A., Walter J.M., Koenig M.;
RA "Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-
RT ocular apraxia 2.";
RL Nat. Genet. 36:225-227(2004).
RN [2]
RP SEQUENCE FROM N.A.

```



RC STRAIN=C57BL/6J;  
RA Morita M.-C., Koenig M.;  
RA Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.  
CC -1- MISCELLANEOUS: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ third party annotation (TRA) entry.  
DR EMBL; BK001523; DAA01946.1;  
SQ SEQUENCE 2646 AA; 297630 MW; CDCFC822443P969 CRC64;  
Query Match 71.1%; Score 40.5; DB 2; Length 2646;  
Beet Local Similarity 66.7%; Pred. No. 36;  
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
Cy 1 RILKCA-SRXAVDXL 14  
Db 1978 RVLVCAPSNAAVDL 1992  
RESULT 11  
SETX\_HUMAN STANDARD; PRT; 2677 AA.  
AC 072333; 075120; 072336; 08MX33; 09H9D1; 09NPV9;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Probable heliase senatxin (EC 3.6.1.-) (SEN1 homolog).  
GN Name=SETX; Synonyms=KIAA0625, SCAR1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, TISSUE SPECIFICITY, VARIANTS  
RP CYS-1152; ASP-1192; GLY-1252 AND ILE-1386, AND VARIANTS A0A2 CYS-305;  
RP TRP-332; LEU-413; SER-1756 AND LEU-2213.  
RX PubMed:14770181; DOI=10.1038/ng1303;  
RA Moreira M.-C., Klur S., Watanabe M., Nemeth A.H., Le Ber I.,  
RA Moniz J.-C., Tranchant C., Aubourg P., Tazir M., Schoels L.,  
RA Pandolfi M., Schulz J.B., Pouget J., Calvas P., Shizuka-Ikeda M.,  
RA Shoji M., Tanaka M., Izatt L., Shaw C.E., W.Zahem A., Dunne E.,  
RA Bonnet P., Benhassine T., Boulsam N., Stevanin G., Brice A.,  
RA Guimaraes J., Mendonca P., Bardot C., Coutinho P., Sequeiros J.,  
RA Duert A., Watter J.-M., Koenig M.;  
RA "Senatxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-  
RT ocular apraxia 2.";  
RL Nat. Genet. 36:225-227(2004).  
[2]  
RP SEQUENCE FROM N.A. (ISOFORM 2), SEQUENCE OF 1762-2677 FROM N.A.  
RP (ISOFORM 1), AND VARIANT VAL-2587.  
RP TISUB=Brain, and Tereocarcinoma;  
RX PubMed:14702039; DOI=10.1038/ng1285;  
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
RA Sekine M., Ohyashiki M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,  
RA Nishiguchi K., Tanai H., Kimeta M., Watanabe M., Hirooka S., Chiba Y.,  
RA Ishida S., Oono Y., Takiguchi S., Watanabe S., Yosida M., Houchi T.,  
RA Kusano Y., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,  
RA Nomura Y., Togiyama S., Komai F., Hara R., Takeuchi K., Arita M.,  
RA Imobe N., Muesehino K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,  
RA Yoshihawa Y., Matsumura H., Ichihara T., Shiohata N., Sano S.,  
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
RA Hishigaki H., Takanebe T., Sugiyama A., Takamoto M., Kawakami B.,  
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
RA Ono T., Yamada K., Fujii Y., Oaki K., Hirao M., Omori Y.,  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,

RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuo Y., Yamashita R.,  
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
RT "Complete sequencing and characterization of 21,243 full-length human  
RT cDNAs.";  
RL Nat. Genet. 36:40-45(2004).  
[3]  
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT ASP-1192.  
RP TISUB=Renal Kidney, and Retina;  
RA Lauber J., Bahr A., Wewes H.-W., Well B., Amid C., Oseanger A.,  
RA Podo G., Han M., Wiemann S.;  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Skuce C.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE OF 15-2677 FROM N.A. (ISOFORM 1), AND VARIANTS GLY-1252 AND  
RP ILE-1386.  
RX TISUB=Brain;  
RX MEDLINE=98403880; PubMed=9734811;  
RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,  
RA Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. X.  
RT The complete sequences of 100 new cDNA clones from brain which can  
RT code for large proteins in vitro.";  
RL DNA Res. 5:165-176(1998).  
RN [6]  
RP SEQUENCE OF 1651-2677 FROM N.A. (ISOFORM 1).  
RP TISUB=Eye;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins P.S., Wagner L., Shemen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrini P., Prange C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huily S.W.,  
RA Villalón D.K., Murray D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Pahay J., Helton E., Keltman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,  
RA Schercher A., Schein J.E., Jones S.J.W., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [7]  
RP VARIANTS ALG4 SER-389 AND HIS-2136, AND TISUB SPECIFICITY.  
RP PubMed:15106121; DOI=10.1086/421054;  
RX Chen Y.-Z., Bennett C.L., Huynh H.M., Blair I.P., Puls I., Irobi J.,  
RA Dietrick I., Abel A., Kemner M.L., Rabin B.A., Nicholson G.A.,  
RA Auer-Grunbach M., Wagner K., De Jonghe P., Griffin J.W.,  
RA Flisbech K.H., Timmerman V., Corbaliath D.R., Chance P.F.;  
RT "DNA/RNA helicase gene mutations in a form of juvenile amyotrophic  
RT lateral sclerosis (ALG4).";  
RL Am. J. Hum. Genet. 74:1128-1135(2004).  
[8]  
RP FUNCTION: Probable helicase, which may be involved in RNA  
RP maturation (by similarity).  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q72333-1; Sequence=D1splayed;  
CC Name=2;  
CC IsoId=Q72333-2; Sequence=VSP\_010532, VSP\_010533;  
CC Note=No experimental confirmation available;

CC -1- TISSUE SPECIFICITY: Highly expressed in skeletal muscle. Expressed  
 CC in heart, placenta, liver and kidney. Weakly expressed in brain  
 CC and lung.  
 CC -1- DISEASE: Defects in SETX are the cause of ataxia-ocular apraxia 2  
 CC (AOA2) [MIM:606002]. AOA2 is an autosomal recessive adolescent-  
 CC onset cerebellar ataxia with additional oculomotor apraxia or with  
 CC associated elevated serum alpha-fetoprotein (AFP), immunoglobulins  
 CC and creatine kinase levels but no oculomotor apraxia. Most  
 CC affected individuals have both oculomotor apraxia and elevated AFP  
 CC levels. The most debilitating feature of the AOA syndrome is the  
 CC progressive neurodegeneration associated with loss of Purkinje  
 CC cells and ectopic location of these cells in the molecular layer.  
 CC -1- DISEASE: Defects in SETX are a cause of amyotrophic lateral  
 CC sclerosis 4 (ALS4) [MIM:602433]. ALS4 is a rare, childhood- or  
 CC adolescent-onset, autosomal dominant form of amyotrophic lateral  
 CC sclerosis that is characterized by slow disease progression, limb  
 CC weakness, severe muscle wasting, and pyramidal signs associated  
 CC with degeneration of motor neurons in the brain and spinal cord.  
 CC Amyotrophic lateral sclerosis denote a heterogeneous group of  
 CC severe, progressive neurological disorders associated with  
 CC degeneration of motor neurons in the cerebral cortex, brain stem,  
 CC and spinal cord. ALS4 includes a long duration of disease, absence  
 CC of overt sensory abnormalities, and the sparing of bulbar and  
 CC respiratory muscles.  
 CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.  
 CC -1- CAUTION: Ref.3 (CAD97857) sequence differs from that shown due to  
 CC a frameshift in position 1626.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AY362728; AAR3367.1; -;  
 CC EMBL; AK001456; BAA91701.1; ALT\_INIT.  
 CC EMBL; AK022902; BAA14299.1; ALT\_INIT.  
 CC EMBL; AK125448; BAC86166.1; -;  
 CC EMBL; BX537849; CAD97857.1; ALT\_FRAME.  
 CC EMBL; BX538166; CAD98045.1; -;  
 CC EMBL; AL159997; CAD13341.1; -;  
 CC EMBL; AL153701; -; NOT\_ANNOTATED\_CDS.  
 CC EMBL; AB014525; BAA31600.2; -;  
 CC EMBL; BC032600; AAB32600.1; ALT\_INIT.  
 CC EMBL; BC032622; AAB32622.1; ALT\_INIT.  
 CC MIM; 608465; -;  
 CC MIM; 606002; -;  
 CC MIM; 602433; -;  
 CC Alternative splicing; ATP-binding; Disease mutation; Helicase;  
 CC Hydrolyase; Neurodegeneration; Nuclear protein; Polymorphism.  
 CC NP\_BIND 1963 1970  
 CC DOMAIN 2070 2087  
 CC FT VASPLIC 625 635  
 CC FT EOMGKTSRKDM -> SLSPNLSYRK (in isoform  
 CC 2).  
 CC FT /FTid=VSP\_010532.  
 CC FT Missing (in isoform 2).  
 CC FT VASPLIC 636 2677  
 CC FT /FTid=VSP\_010533.  
 CC FT T -> I (in ALS4; heterozygous).  
 CC FT /FTid=VAR\_018776.  
 CC FT W -> C (in AOA2; homozygous in an  
 CC Algerian family).  
 CC FT /FTid=VAR\_018777.  
 CC FT R -> W (in AOA2; in a French family).  
 CC FT /FTid=VAR\_018778.  
 CC FT L -> S (in ALS4; heterozygous).  
 CC FT /FTid=VAR\_018779.  
 CC FT P -> L (in AOA2; homozygous in a Tunisian  
 CC family).  
 CC FT /FTid=VAR\_018780.  
 CC FT A -> G (in dbSNP:882709).  
 CC FT VARIANT 660 660

Query Match 71.1%; Score 40.5; DB 1; Length 2677;  
 Best Local Similarity 66.7%; Pred. No. 36;  
 Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
 Oy 1 RILXCA-SNXAVDXL 14  
 Db 2002 RVLVCAPSNNAVDL 2016  
 RESULT 12  
 ID 06CW68 PRELIMINARY; PRT; 969 AA.  
 AC 06CW68;  
 DT 25-OCT-2004 (TREMblrel. 28, Created)  
 DT 25-OCT-2004 (TREMblrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMblrel. 28, Last annotation update)  
 DE Kluveromyces lactis strain NRRL Y-1140 chromosome B of strain NRRL Y-  
 DE 1140 of Kluveromyces lactis.  
 GN ORENames=KLA0B064359;  
 OS Kluveromyces lactis NRRL Y-1140.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Kluveromyces.  
 OX NCBI\_TaxID=2845590;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL Y-1140;  
 RG Genolevures;  
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S., Yalla E.,  
 RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Yalla E.,  
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Barray S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boissarie A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
 RA Hantaye F., Hémeguin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicard J.M., Nikolaki M., Ozias S., Ozier-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Sureau A.,  
 RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
 RA Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Galliardin C., Weissenbach J.,  
 RA Wincker P., Souciet J.L.,  
 RT "Genome evolution in yeasts."  
 RL Nature 430:35-44(2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL Y-1140;  
 RG Genoscope;  
 RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR382122; CAH02214.1; -;  
 DR InterPro: IPR006935; Resili.  
 DR InterPro: IPR000504; RNA\_rec\_mot.  
 DR Pfam; PF04851; Resili1.1; UNKNOWN 1.  
 DR PROSITE; PS00030; RRM\_NP\_1; 208C14P1B6BCE29 CRC64;  
 SQ SEQUENCE 969 AA; 108943 MW; 208C14P1B6BCE29 CRC64;  
 Query Match 69.3%; Score 39.5; DB 2; Length 969;  
 Best Local Similarity 66.7%; Pred. No. 22;  
 Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
 Oy 1 RILXCA-SNXAVDXL 14  
 Db 452 RVLVCAPSNNAVDL 466  
 RESULT 13  
 ID 07SDS7 PRELIMINARY; PRT; 1000 AA.  
 AC 07SDS7;  
 DT 05-JUL-2004 (TREMblrel. 27, Created)  
 DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMblrel. 27, Last annotation update)  
 DE ABR022Cp.

```

GN OREName=ABR022C;
OS Aeshya gossypii (Yeast) (Bremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Bremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10895;
RA Voegel S.E., Dietrich F.S., Brachat S., Lerch A., Gaffney T.,
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB016815; AAS50792.2; -.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_Arpase.
DR SMART; SM00382; AAA; 1.
KM ATP-binding.
SQ SEQUENCE 1000 AA; 11963 MW; 39D4257F675B72A6 CRC64;

Query Match 69.3%; Score 39.5; DB 2; Length 1000;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 1 RILKCA-SNXAVDXL 14
Db 463 RVLVCASNTAVDHL 477

RESULT 14
Q88B60 PRELIMINARY; PRT; 191 AA.
ID Q88B60;
AC Q88B60;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE RNA polymerase sigma-70 factor, ECF subfamily.
OS Pseudomonas putida (strain KT2440).
GN Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160486;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLIN=22423060; PubMed=12534463;
RA Nelson K.E., Wetzel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Pouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.W., Deboy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Morzzer A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hohelsel J., Stratz M., Helm S.,
RA Kiewitz C., Bisen J.A., Timmis K.N., Duesternoeft A., Tuenmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AB016791; AAN70181.1; -.
DR TIGR; PP4608; -.
DR GO; GO:0016987; P:sigma factor activity; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006352; P:transcription initiation; IEA.
DR InterPro; IPR007627; Sigma70_r2.
DR InterPro; IPR007630; Sigma70_r4.
DR Pfam; PF04542; Sigma70_r2; 1.
DR Pfam; PF04545; Sigma70_r4; 1.
KM Complete proteome.
SQ SEQUENCE 191 AA; 21313 MW; 505CB7619978B999 CRC64;

Query Match 68.4%; Score 39; DB 2; Length 191;
Best Local Similarity 53.8%; Pred. No. 5.6;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ILKCSNXAVDXL 14

```

```

Db 72 ILKCAIADLH 84

RESULT 15
Q7PP43 PRELIMINARY; PRT; 611 AA.
ID Q7PP43;
AC Q7PP43;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE ENSANGP000001315 (fragment).
GN Name=ENSANG00000011026;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anophelae.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008960; EAA11070.2; -.
FT NON_TER 1 611
FT NON_TER 1 611
SQ SEQUENCE 611 AA; 68818 MW; 417B9DB5546DE23P CRC64;

Query Match 67.5%; Score 38.5; DB 2; Length 611;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 1 RILKCA-SNXAVDXL 14
Db 202 RVLVCASNTAVDVI 216

```

Search completed: April 18, 2005, 08:15:39  
 Job time : 61.4526 secs

(folded)  
this page blank (uspto)



PI Peltz S, Czaplinski K, Dimman JD;  
 XX WPI; 2003-810549/76.  
 XX  
 PT Identifying an agent that increases nonsense suppression, for antiviral  
 PT therapy, by contacting modulator of translation termination (Mtt1) in  
 PT *Saccharomyces cerevisiae* with a test agent, and detecting specific  
 PT binding to Mtt1.  
 XX  
 PS Disclosure: Col 43-44; 0pp; English.  
 XX  
 CC The invention relates to a method of identifying an agent that increases  
 CC nonsense suppression, by contacting modulator of translation termination  
 CC (Mtt1) also referred to as helicase B (HCSB) in *Saccharomyces cerevisiae*.  
 CC The method is useful for identifying compositions or agents which  
 CC increase nonsense suppression. The invention may also be used for  
 CC antiviral therapy and for suppression of pathological nonsense mutations.  
 CC The present sequence is *Saccharomyces cerevisiae* motif peptide  
 CC  
 SO Sequence 17 AA;  
 Query Match 74.1%; Score 40; DB 7; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 0.0046;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VVIDEXXQAXXXXXXIIPI 17  
 DB 1 VVIDEXXQAXXXXXXIIPI 17  
 RESULT 2  
 ADP44105  
 ID ADP44105 standard; peptide; 17 AA.  
 AC  
 XX ADP44105;  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Yeast translation termination modulation protein motif III.  
 XX  
 KW gene therapy; translation termination; RNA helicase; MTT1;  
 KW frameshift frequency; aberrant transcript degradation;  
 KW peptidyl transferase modulation; beta-thalassemia; beta-globin;  
 KW Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
 KW Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;  
 KW Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic fibrosis;  
 KW Kidney Stone; Familial hypercholesterolemia; Retinitis Pigmentosa;  
 KW Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; Yeast.  
 KW  
 XX *Saccharomyces cerevisiae*.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 6..7  
 FT /note= "Any amino acid"  
 FT Misc-difference 10..14  
 FT /note= "Any amino acid"  
 XX  
 XX US2004115787-A1.  
 XX  
 XX 17-JUN-2004.  
 XX  
 XX 28-AUG-2003; 2003US-00652334.  
 XX  
 XX 22-JUL-1998; 98US-0093685P.  
 XX  
 XX 22-JUL-1999; 99US-00359268.  
 XX  
 XX (PELTZ) PELTZ S.  
 XX (CZAP/) CZAPLINSKI K.  
 XX (DIMM/) DIMMAN J D.  
 XX  
 XX Peltz S, Czaplinski K, Dimman JD;

DR WPI; 2004-449400/42.  
 XX  
 PT Identifying a test composition or agent that modulates the efficiency of  
 PT translation termination comprises contacting the MTT1 with the test  
 PT composition or agent, and determining if the test composition or agent  
 PT inhibits the MTT1.  
 XX  
 PS Claim 35; SEQ ID NO 3; 41pp; English.  
 XX  
 CC The invention relates to a method of identifying a test composition that  
 CC modulates the efficiency of translation termination comprising contacting  
 CC the RNA helicase MTT1 with a composition or agent under conditions  
 CC permitting binding between the MTT1 and the composition, detecting  
 CC specific binding of the test composition or agent to the MTT1, and  
 CC determining if the test composition or agent inhibits the MTT1. The  
 CC composition and methods are useful for modulating the fidelity of  
 CC translation termination or for identifying agents that affect the  
 CC functional activity of mRNAs by altering frameshift frequency, permit  
 CC monitoring of a termination event, promote degradation of aberrant  
 CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl  
 CC transferase activity during initiation, elongation, termination and mRNA  
 CC degradation of translation. The agents, which may be antagonists or  
 CC agonists, are useful in screening, diagnostic and therapeutic purposes,  
 CC for diseases or conditions resulting from or cause premature translation,  
 CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular  
 CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,  
 CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,  
 CC Hirschsprung disease, Cystic fibrosis, Kidney Stones, Familial  
 CC hypercholesterolemia, Retinitis Pigmentosa, or Neurofibromatosis,  
 CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents  
 CC the amino acid sequence of the yeast translation termination modulation  
 CC protein motif III.  
 CC  
 SO Sequence 17 AA;  
 Query Match 74.1%; Score 40; DB 8; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 0.0046;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VVIDEXXQAXXXXXXIIPI 17  
 DB 1 VVIDEXXQAXXXXXXIIPI 17  
 RESULT 3  
 AA77806  
 ID AA77806 standard; peptide; 18 AA.  
 XX  
 XX AA77806;  
 DT 31-MAY-2000 (first entry)  
 XX  
 DE Motif III comprised in a gene modulating translation termination.  
 XX  
 KW Helicase B; HCSB; MTT1; modulator of translation termination; eRF1; eRF3;  
 KW eukaryotic release factor; peptidyl transferase; beta-thalassemia;  
 KW beta-globin; Duchenne/Becker Muscular Dystrophy; antianemic.  
 KW  
 XX Unidentified.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 1..18  
 FT /note= "residues indicated Xaa are unspecified"  
 FT  
 XX  
 XX WO200005586-A2.  
 XX  
 XX 03-FEB-2000.  
 XX  
 XX 22-JUL-1999; 99WO-US016802.  
 XX  
 XX 22-JUL-1998; 98US-00120435.  
 XX  
 XX (UYNE-) UNITV NEW JERSEY.

PI Peltz S, Czaplinski K, Dimman JD;  
 XX  
 XX WPI; 2000-171458/15.  
 DR  
 PR New multiprotein complex which can modulate peptidyl transferase activity  
 PR during translation, useful to treat diseases associated with peptidyl  
 PR transferase activity e.g. Duchenne/Becker Muscular Dystrophy.  
 PT  
 PS Claim 35; Page 79; 89pp; English.  
 XX  
 XX The invention provides a new multiprotein complex which can modulate  
 CC peptidyl transferase activity during translation. The complex comprises  
 CC the gene encoding Helicase B (HCSB, renamed MTI), for Modulator of  
 CC Translation Termination) and the conserved proteins known to interact and  
 CC carry out translation termination in eukaryotic cells, peptidyl-  
 CC eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to  
 CC modulate peptidyl transferase activity during translation in a cell. It  
 CC can be administered therapeutically combined with a carrier in  
 CC pharmaceutical compositions to treat diseases associated with peptidyl  
 CC transferase activity, especially diseases resulting from a nonsense or  
 CC frameshift mutation e.g. beta-thalassemia, beta-globin, Duchenne/Becker  
 CC Muscular Dystrophy etc. It can be used to identify disease conditions  
 CC involving a defect in the complex, by transfecting cells with encoding  
 CC nucleic acid and determining the proportion of defective complex before  
 CC and after transfection. It is also useful to screen for drugs involved in  
 CC peptidyl transferase activity during translation, inhibiting the  
 CC interaction between MTI and eRF3 or involved in enhancing translation  
 CC termination. Vectors comprising polynucleotides encoding the complex (or  
 CC antisense sequences) can be constructed and introduced into cells to  
 CC interfere with complex expression and so modulate the efficiency of  
 CC translation termination of mRNA and/or degradation of aberrant  
 CC transcripts in a cell. Agents binding to the complex can be identified  
 CC and included in therapeutic compositions useful as above, and/or used to  
 CC modulate peptidyl transferase activity during translation in cells. They  
 CC are also useful to modulate the efficiency of translation termination of  
 CC mRNA at a nonsense codon and/or promote degradation of aberrant  
 CC transcripts in cells. The method can be used to identify agents/  
 CC compositions modulating binding to MTI, useful to identify genes.  
 CC Sequences AAY77804-812 represent motifs I-IX comprised in the genes of  
 CC interest, used for modulating translation termination  
 CC  
 SO Sequence 18 AA;  
 Query Match 74.1%; Score 40; DB 3; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 0.0049;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
 Oy 1 VVIDEXXQAXXXXXXPI 17  
 |||||  
 Db 1 VVIDEXXQAXXXXXXPI 17  
 |||||  
 RESULT 4  
 ID ADN20231 standard; protein; 650 AA.  
 AC ADN20231;  
 XX  
 XX 02-DEC-2004 (first entry)  
 DT  
 XX  
 XX Bacterial polypeptide #2884.  
 DB  
 XX Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.  
 XX  
 XX Bacteria.  
 OS

```

PN  US2003233675-A1.
XX
XX  18-DEC-2003.
PD
XX
XX  20-FEB-2003; 2003US-00369493.
PF
XX
XX  21-FEB-2002; 2002US-0360039P.
PR
XX
XX  (CAOY/) CAO Y.
PA  (HINK/) HINKLE G J.
PA  (SLAT/) SLATER S C.
PA  (CHEN/) CHEN X.
PA  (GOLD/) GOLDMAN B S.
PI
PI  Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
DR  WPI; 2004-061375/06.
XX
XX  New recombinant DNA construct comprising a promoter positioned to provide
PT  for expression of a polynucleotide encoding a polypeptide from a
PT  microbial source, useful for producing plants with improved properties.
XX
XX  Claim 1; SEQ ID NO 2884; 122pp; English.
XX
XX  The invention relates to a recombinant DNA construct comprising a
CC  promoter functional in a plant cell, where the promoter is positioned to
CC  provide for expression of a polynucleotide encoding a polypeptide from a
CC  microbial source. The invention also relates to a transformed plant
CC  comprising the recombinant DNA construct and a method of producing a
CC  transformed plant having an improved property. The plant is a crop plant
CC  such as maize or soybean. The method of producing a transformed plant
CC  having an improved property comprises transforming a plant with the
CC  recombinant DNA construct and growing the transformed plant, where the
CC  polynucleotide or polypeptide is useful for improving plant properties.
CC  The recombinant DNA construct is useful for producing plants with
CC  improved plant properties, e.g. improved cold, heat or drought tolerance,
CC  tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC  increased resistance to plant disease, better growth rate by modification
CC  of the cell cycle pathway with plant growth regulators, increased rate of
CC  homologous recombination, modified seed oil or protein yield and/or
CC  content, improved yield by modification of carbohydrate, nitrogen or
CC  phosphorus use and/or uptake, by modification of photosynthesis or by
CC  providing improved plant growth and development under at least one stress
CC  condition, improved lignin production or improved galactomanan
CC  production. This sequence represents a bacterial polypeptide used in the
CC  scope of the invention. Note: The sequence data for this patent did not
CC  form part of the printed specification but was obtained in electronic
CC  format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX  Sequence 650 AA;
SQ
XX
XX  Query Match 60.2%; Score 32.5; DB 8; Length 650;
XX  Best Local Similarity 50.0%; Pred. No. 26;
XX  Matches 9; Conservative 1; Mismatches 7; Indels 1; Gaps 1
XX
XX  1 VVIDEXXQAXX-XXXIP1 17
XX  ||:|||||
XX  ||:|||||
XX  380 VVVDENASQATIPSIIP1 397
XX
XX  RESULT 5
XX  ABP26966
XX  ID ABP26966 standard; protein, 318 AA.
XX
XX  ABP26966;
XX
XX  02-JUL-2002 (first entry)
XX
XX  Streptococcus polypeptide SEQ ID NO 3108.
XX
XX  Streptococcus polypeptide SEQ ID NO 3108.
XX
XX  Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX  group A streptococcus; Streptococcus pyogenes; antibacterial;
XX  anti-inflammatory; infection; vaccine; meningitis; gene therapy.

```

XX OS Streptococcus agalactiae.  
XX PN WO200234771-A2.  
XX PD 02-MAY-2002.  
XX PF 29-OCT-2001; 2001WO-GB004789.  
XX PR 27-OCT-2000; 2000GB-00026333.  
XX PR 24-NOV-2000; 2000GB-00028727.  
XX PR 07-MAR-2001; 2001GB-00005640.  
XX PA (CHIR-) CHIRON SPA.  
XX PA (GENO-) INST GENOMIC RES.  
XX PI Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C,  
XX PI Tetteijn H;  
XX DR MPI; 2002-352536/38.  
XX DR N-PSDB; ABN67597.  
XX PT New Streptococcus protein for the treatment or prevention of infection or  
XX PT disease caused by Streptococcus bacteria, such as meningitis, and for  
XX PT detecting a compound that binds to the protein.  
XX PS Claim 1; Page 3464; 4525pp; English.  
XX CC The invention relates to a protein (ABP25413-ABP30895) from group B  
XX CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
XX CC the specification. The proteins have antibacterial and anti-inflammatory  
XX CC activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and  
XX CC antibodies that bind (1) are used in the manufacture of medicaments for  
XX CC the treatment or prevention of infection or disease caused by  
XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
XX CC Nucleic acids encoding (1) are used to detect Streptococcus in a  
XX CC biological sample. (1) is used to determine whether a compound binds to  
XX CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be  
XX CC used as a vaccine or diagnostic composition. The disease caused by  
XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
XX CC acid encoding (1) may be used to recombinantly produce (1) and may be  
XX CC used in gene therapy. Antibodies to (1) are used for affinity  
XX CC chromatography, immunoassays, and distinguishing/identifying  
XX CC Streptococcus proteins  
XX SQ Sequence 318 AA;  
QY Query Match 59.3%; Score 32; DB 5; Length 318;  
Best Local Similarity 41.2%; Pred. No. 16;  
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;  
Db 1 VVIDEXXQAXXXXXXIP1 17  
117 VIVDEKQDAFTNPTKPI 133  
RESULT 6  
ABU34245  
ID ABU34245 standard; protein; 710 AA.  
XX AC ABU34245;  
XX DT 19-JUN-2003 (first entry)  
XX DE Protein encoded by Prokaryotic essential gene #19772.  
XX KM Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX OS Mycobacterium avium.  
XX AC WO200277183-A2.  
XX PN

PD 03-OCT-2002.  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX PR 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362699P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,  
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX DR MPI; 2003-029926/02.  
XX DR N-PSDB; ACA38115.  
XX PT New antisense nucleic acids, useful for identifying proteins or screening  
XX PT for homologous nucleic acids required for cellular proliferation to  
XX PT isolate candidate molecules for rational drug discovery programs.  
XX PS Claim 25; SEQ ID NO 62169; 1766pp; English.  
XX CC The invention relates to an isolated nucleic acid comprising any one of  
XX CC the 6213 antisense sequences given in the specification where expression  
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid  
XX CC encoding a polypeptide whose expression is inhibited by the antisense  
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX CC polypeptide or its fragment whose expression is inhibited by the  
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding  
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX CC proliferation or the activity of a gene in an operon required for  
XX CC proliferation; (7) identifying a compound that influences the activity of  
XX CC the gene product or that has an activity against a biological pathway  
XX CC required for proliferation, or that inhibits cellular proliferation; (8)  
XX CC identifying a gene required for cellular proliferation or the biological  
XX CC pathway in which a proliferation-required gene or its gene product lies  
XX CC or a gene on which the test compound that inhibits proliferation of an  
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX CC compound's activity; (11) a culture comprising strains in which the gene  
XX CC product is overexpressed or underexpressed; (12) determining the extent  
XX CC to which each of the strains is present in a culture or collection of  
XX CC strains; or (13) identifying the target of a compound that inhibits the  
XX CC proliferation of an organism. The antisense nucleic acids are useful for  
XX CC identifying proteins or screening for homologous nucleic acids required  
XX CC for cellular proliferation to isolate candidate molecules for rational  
XX CC drug discovery programs, or for screening homologous nucleic acids  
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,  
XX CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of  
XX CC the target prokaryotic essential genes. Note: The sequence data for this  
XX CC patent did not form part of the printed specification, but was obtained  
XX CC in electronic format directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 710 AA;  
QY Query Match 59.3%; Score 32; DB 6; Length 710;  
Best Local Similarity 31.2%; Pred. No. 40;  
Matches 5; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
Db 2 VVIDEXXQAXXXXXXIP1 17  
558 IMDETVOAAAADGAPL 573  
RESULT 7  
ADN19868  
ID ADN19868 standard; protein; 1944 AA.  
XX AC ADN19868;  
XX PN



DT 02-DEC-2004 (first entry)  
 XX Bacterial polypeptide #2521.  
 XX  
 KW Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.  
 XX  
 OS Bacteria.  
 XX  
 XX US2003233675-A1.  
 XX  
 PD 18-DEC-2003.  
 XX  
 PP 20-FEB-2003; 2003US-00369493.  
 XX  
 PR 21-FEB-2002; 2002US-0360039P.  
 XX  
 PA (CAOY/) CAO Y.  
 PA (HINK/) HINKLE G J.  
 PA (SLAT/) SLATER S C.  
 PA (CHEN/) CHEN X.  
 PA (GOLD/) GOLDMAN B S.  
 XX  
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
 DR WPI; 2004-061375/06.  
 XX  
 PT New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.  
 XX  
 PS Claim 1; SEQ ID NO 2521; 122pp; English.  
 XX  
 CC The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition. Improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at seqdata.uspto.gov/sequence.html.  
 XX  
 SQ Sequence 1944 AA;  
 Query Match 59.3%; Score 32; DB 8; Length 1944;  
 Best Local Similarity 46.7%; Pred. NO. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
 GY 1 VVIDEXXQAXXXXXX 15  
 DB 1513 VVIDEAAQAVELSSI 1527

RESULT 8  
 ADS44332  
 ID ADS44332 standard; protein; 712 AA.  
 XX  
 AC ADS44332;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE Bacterial polypeptide #22762.  
 XX  
 KW Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.  
 XX  
 OS Bacteria.  
 XX  
 XX US2003233675-A1.  
 XX  
 PD 18-DEC-2003.  
 XX  
 PP 20-FEB-2003; 2003US-00369493.  
 XX  
 PR 21-FEB-2002; 2002US-0360039P.  
 XX  
 PA (CAOY/) CAO Y.  
 PA (HINK/) HINKLE G J.  
 PA (SLAT/) SLATER S C.  
 PA (CHEN/) CHEN X.  
 PA (GOLD/) GOLDMAN B S.  
 XX  
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
 DR WPI; 2004-061375/06.  
 XX  
 PT New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.  
 XX  
 PS Claim 1; SEQ ID NO 22762; 122pp; English.  
 XX  
 CC The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition. Improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at seqdata.uspto.gov/sequence.html.  
 XX  
 SQ Sequence 712 AA;  
 Query Match 57.4%; Score 31; DB 8; Length 712;

Best Local Similarity 66.7%; Pred. No. 72;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
|:|||||  
Db 438 VVIDEASQA 446

## RESULT 9

ID ADR68058 standard; protein; 818 AA.  
XX ADR68058;

AC ADR68058;

DT 06-MAY-2004 (first entry)

XX Female reproductive cell formation magatama mutant (MAA3) protein.

XX magatama mutant; MAA3; female reproductive cell formation; plant; seed;  
KW transgenic; foodstuff; endosperm.

XX Arabidopsis thaliana.

XX JP2004024106-A.

PD 29-JAN-2004.

XX 25-JUN-2002; 2002JP-00185184.

PR 25-JUN-2002; 2002JP-00185184.

PA (KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.

XX WPI: 2004-127085/13.

DR N-PSDB; ADR68059.

PT Novel magatama mutant protein derived from Arabidopsis thaliana for  
XX female reproductive cell formation.

PS Claim 1; SEQ ID NO 1; 40pp; Japanese.

XX The invention relates to a novel magatama mutant (MAA3) protein for  
CC female reproductive cell formation, having an 818 amino acid sequence,  
CC given in the specification or a sequence which has deletion, substitution  
CC or addition in the amino acid(s) of the 818 amino acid protein. The  
CC invention further relates to: a gene encoding the MAA3 protein; DNA  
CC containing the MAA3 gene, as a 2706 nucleotide sequence, given in the  
CC specification or a sequence which hybridizes under stringent condition  
CC with the 2706bp nucleic acid; a vector containing the MAA3 gene or one  
CC part of the gene; a transformed cell transformed by the vector; a  
CC nucleotide containing the partial sequence of continuous 14 bases or more  
CC in the 2706bp nucleic acid or its complementary sequence; and a  
CC transforming plant with characters modified by making a female  
CC reproductive cell lethality or forming only the seed which does not  
CC contain a specific allele. The MAA3 gene is derived from Arabidopsis  
CC thaliana and is useful for making a female reproductive cell a lethality  
CC by suppressing the expression of the MAA3 gene. The MAA3 gene is useful  
CC for setting the number of seed to 1 for 2n by inserting antisense MAA3  
CC gene in places of plant genome. The novel method of the invention is  
CC useful for producing transgenic plant-derived foodstuffs, which contain a  
CC foreign gene in neither a seed nor an endosperm. The MAA3 nucleotide  
CC fragment is useful as probe for detecting the MAA3 gene and is also  
CC useful for modulating the expression of the MAA3 gene. This sequence  
CC represents the 818 amino acid MAA3 protein of the invention.

XX Sequence 818 AA;

QY Query Match 57.4%; Score 31; DB 8; Length 818;

Best Local Similarity 66.7%; Pred. No. 85;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
|:|||||

Db 487 VVIDEAAQA 495

## RESULT 10

ID ADN21359 standard; protein; 830 AA.  
XX ADN21359;

AC ADN21359;

DT 02-DEC-2004 (first entry)

XX Bacterial polypeptide #4012.

XX Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.

XX Bacteria.

XX US2003233675-A1.

PD 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

PR 21-FEB-2002; 2002US-0360039P.

PA (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI: 2004-061375/06.

PT New recombinant DNA construct comprising a promoter positioned to provide  
XX for expression of a polynucleotide encoding a polypeptide from a  
XX microbial source, useful for producing plants with improved properties.

PS Claim 1; SEQ ID NO 4012; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX Sequence 830 AA;

Query Match 57.4%; Score 31; DB 8; Length 830;  
 Best Local Similarity 66.7%; Pred. No. 86;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
 |||||  
 DB 477 VVIDEXXQA 485

## RESULT 11

ABO79216 standard; protein; 302 AA.

ABO79216;

29-JUL-2004 (first entry)

Pseudomonas aeruginosa polypeptide #11391.

Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

Pseudomonas aeruginosa.

US6551795-B1.

22-APR-2003.

18-FEB-1999; 99US-00252991.

18-FEB-1998; 98US-0074788P.

27-JUL-1998; 98US-0094190P.

(GENO-) GENOME THERAPEUTICS CORP.

Rubenfeld MJ, Nolling J, Deloughery C, Bush D;

WPI; 2003-615309/58.

N-PSDB; ABD12787.

Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.

Disclosure; SEQ ID NO 27962; 455bp; English.

The invention relates to Pseudomonas aeruginosa polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a P. aeruginosa nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-P. aeruginosa drugs, as templates for recombinant production of P. aeruginosa-derived peptides or polypeptides, as target components for diagnosis and/or treatment of P. aeruginosa-caused infection, and in detection of P. aeruginosa sequences or other sequences of Pseudomonas species using biochip technology. Sequences ABO67826-ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html)

Sequence 302 AA;

Query Match 55.6%; Score 30; DB 7; Length 302;

Best Local Similarity 77.8%; Pred. No. 50;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
 |||||  
 DB 70 VVIDEXXQA 78

RESULT 12  
 ABO7940 standard; protein; 352 AA.  
 ID ABO7940

ABO7940;

20-NOV-2003 (first entry)

Allotococcus otitis antigenic protein SEQ ID NO:1880.

Allotococcus otitidis; antigenic protein; immunogenic; immunisation; gene therapy; Gram-positive bacterium; infection.

Allotococcus otitis.

WO2003048304-A2.

12-JUN-2003.

25-NOV-2002; 2002WO-US036123.

29-NOV-2001; 2001US-0333777P.

18-NOV-2002; 2002US-0426742P.

(AMHP) WYETH HOLDINGS CORP.

Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;

WPI; 2003-505284/47.

N-PSDB; ABO7939.

New Allotococcus otitidis polynucleotides and polypeptides, useful for treating and diagnosing diseases, drug screening assays and monitoring of effects during drug clinical trials.

Claim 33; SEQ ID NO 1880; 1019bp; English.

The present invention describes an isolated polynucleotide (1) of Allotococcus otitidis genomic DNA, which encodes an antigenic protein. Allotococcus otitidis is a Gram-positive bacterium. Also described: (1) an isolated polypeptide that is encoded by the polynucleotide (1); (2) an expression vector comprising the novel isolated polynucleotide (1), its complement, degenerate variant or fragment; (3) a genetically engineered host cell, transfected, transformed or infected with the vector of (2); (4) an antibody specific for the polypeptide of (1); (5) an immunogenic composition comprising the polypeptide, its complement, biological equivalent or fragment, or the polynucleotide that is comprised in the expression vector; (6) a pharmaceutical composition comprising the polypeptide of (1) and a carrier; (7) a protein chip comprising an array of the polypeptides of (1), their biological equivalent or fragment; (8) immunising against Allotococcus otitidis by administering to a host the otitidis in the biological sample; (9) detecting and/or identifying Allotococcus otitidis in the novel polynucleotide, its degenerate variant or fragment, containing the novel polynucleotide, its degenerate variant or fragment, or the antibody of (4); and (11) producing a polypeptide by culturing the genetically engineered host cell under conditions suitable to produce the polypeptide from the culture. (1) can be used in gene therapy. The polynucleotides, polypeptides, antibodies and compositions of the present invention can be used for treating and diagnosing diseases, drug screening assays and monitoring of effects during drug clinical trials. The polynucleotides are useful for expressing and detecting Allotococcus otitidis. The present sequence represents an Allotococcus otitidis antigen protein from the present invention.

Sequence 352 AA;

Query Match 55.6%; Score 30; DB 6; Length 352;

Best Local Similarity 40.0%; Pred. No. 60;  
 Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 IDEXXQAXXXXXXPI 17  
 :|||  
 :|||



Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;  
 QY 3 IDEXXQAXXXKXIP1 17  
 :||| :|||  
 Db 133 LDELNQASADLGLPI 147

## RESULT 15

ADB07946  
 ID ADB07946 standard; protein; 388 AA.

ADB07946;  
 AC

DT 20-NOV-2003 (first entry)

DE Alloiococcus ostitis antigenic protein SEQ ID NO:1886.

KW Alloiococcus ostitidis; antigenic protein; immunogenic; immunisation;  
 gene therapy; Gram-positive bacterium; infection.

OS Alloiococcus ostitis.

PN WO2003048304-A2.

PD 12-JUN-2003.

PP 25-NOV-2002; 2002WO-US036123.

PR 29-NOV-2001; 2001US-033777P.

PR 18-NOV-2002; 2002US-0426742P.

PA (AMHP ) WYETH HOLDINGS CORP.

PI Fletcher LD, Mcmichael JC, Russell DP, Zagursky RJ;

PI MPI; 2003-505284/47.

DR N-PSDB; ADB07945.

FT New Alloiococcus ostitidis polynucleotides and polypeptides, useful for  
 treating and diagnosing diseases, drug screening assays and monitoring of  
 effects during drug clinical trials.

PS Claim 33; SEQ ID NO 1886; 1019pp; English.

XX The present invention describes an isolated polynucleotide (1) of

CC Alloiococcus ostitidis genomic DNA, which encodes an antigenic protein.

CC Alloiococcus ostitidis is a Gram-positive bacterium. Also described: (1)

CC an isolated polypeptide that is encoded by the polynucleotide (1); (2) an

CC expression vector comprising the novel isolated polynucleotide (1), its

CC complement, degenerate variant or fragment; (3) a genetically engineered

CC host cell, transfected, transformed or infected with the vector of (2);

CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic

CC composition comprising the polypeptide, its complement, biological

CC equivalent or fragment, or the polynucleotide that is comprised in the

CC expression vector; (6) a pharmaceutical composition comprising the

CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array

CC of the polypeptides of (1), their biological equivalent or fragment; (8)

CC immunising against Alloiococcus ostitidis by administering to a host the

CC immunogenic composition; (9) detecting and/or identifying Alloiococcus

CC ostitidis in the biological sample; (10) a kit comprising a container

CC containing the novel polynucleotide, its degenerate variant or fragment,

CC or the antibody of (4); and (11) producing a polypeptide by culturing the

CC genetically engineered host cell under conditions suitable to produce the

CC polypeptide from the culture. (1) can be used in gene therapy. The

CC polynucleotides, polypeptides, antibodies and compositions of the present

CC invention can be used for treating and diagnosing diseases, drug

CC screening assays and monitoring of effects during drug clinical trials.

CC The polynucleotides are useful for expressing and detecting Alloiococcus

CC ostitidis. The present sequence represents an Alloiococcus ostitidis

CC antigen protein from the present invention.

CC Sequence 388 AA;

CC SQ

Query Match 55.6%; Score 30; DB 6; Length 388;  
 Best Local Similarity 40.0%; Pred. No. 67;  
 Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;  
 QY 3 IDEXXQAXXXKXIP1 17  
 :||| :|||  
 Db 136 LDELNQASADLGLPI 150

Search completed: April 18, 2005, 08:03:44  
 Job time : 85.7389 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model1

Run on: April 18, 2005, 07:37:21 ; Search time 21.219 Seconds  
(without alignments)  
59.807 Million cell updates/sec

Title: US-10-652-334-3

Sequence: 1 VVIDEXXQAXXXXXXIP1 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCUTS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	74.1	17	US-09-359-268A-3	Sequence 3, Appl1
2	31	57.4	513	US-09-902-540-13564	Sequence 13564, A
3	30	55.6	302	US-09-252-991A-17962	Sequence 27962, A
4	30	55.6	993	US-09-538-092-1100	Sequence 1100, Ap
5	29	53.7	773	US-09-248-796A-20994	Sequence 20994, A
6	29	53.7	161	US-09-252-991A-17847	Sequence 17847, A
7	29	53.7	380	US-09-359-268A-29	Sequence 2529, Ap
8	29	53.7	413	US-09-540-236-2529	Sequence 11693, A
9	29	53.7	671	US-09-902-540-11693	Sequence 29, Appl
10	29	53.7	971	US-08-724-354D-22	Sequence 22, Appl
11	29	53.7	971	US-09-270-984A-22	Sequence 8, Appl1
12	29	53.7	971	US-09-177-431-8	Sequence 19347, A
13	28.5	52.8	917	US-09-248-796A-19347	Sequence 9, Appl1
14	28	51.9	30	US-08-724-354D-9	Sequence 23748, A
15	28	51.9	30	US-09-270-984A-9	Sequence 6445, Ap
16	28	51.9	200	US-09-252-991A-23748	Sequence 7277, Ap
17	28	51.9	305	US-09-328-352-6445	Sequence 27, Appl
18	28	51.9	339	US-09-489-038A-1277	Sequence 4, Appl1
19	28	51.9	377	US-09-910-430-27	Sequence 4, Appl1
20	28	51.9	1043	US-08-724-354D-4	Sequence 2, Appl1
21	28	51.9	1043	US-09-270-984A-4	Sequence 10116, A
22	28	51.9	1118	US-08-724-354D-2	Sequence 16, Appl
23	28	51.9	1118	US-09-270-984A-2	Sequence 34183, A
24	28	51.9	1140	US-09-949-016-10116	Sequence 49400, A
25	27	50.0	11	US-09-359-268A-16	
26	27	50.0	58	US-09-270-767-34183	
27	27	50.0	58	US-09-270-767-49400	

28	27	50.0	154	4	US-09-723-830-2	Sequence 2, Appl1
29	27	50.0	211	4	US-09-949-016-9913	Sequence 9913, Ap
30	27	50.0	400	4	US-09-198-452A-538	Sequence 538, App
31	27	50.0	400	4	US-09-438-185A-501	Sequence 501, App
32	27	50.0	414	4	US-09-359-268A-28	Sequence 28, Appl
33	27	50.0	457	1	US-08-206-006-2	Sequence 2, Appl1
34	27	50.0	472	4	US-09-359-268A-26	Sequence 26, Appl
35	27	50.0	503	4	US-09-949-016-6578	Sequence 6578, Ap
36	27	50.0	683	4	US-09-538-092-483	Sequence 483, App
37	26	48.1	47	4	US-09-471-276-1388	Sequence 1388, Ap
38	26	48.1	125	4	US-09-107-532A-6413	Sequence 6413, Ap
39	26	48.1	149	4	US-09-270-767-33094	Sequence 33094, A
40	26	48.1	149	4	US-09-270-767-48311	Sequence 48311, A
41	26	48.1	176	4	US-09-489-039A-7716	Sequence 7716, Ap
42	26	48.1	219	1	US-08-166-529-4	Sequence 4, Appl1
43	26	48.1	219	1	US-08-640-386A-4	Sequence 4, Appl1
44	26	48.1	219	2	US-08-751-767A-2	Sequence 18, Appl
45	26	48.1	219	4	US-09-462-941-18	

## ALIGNMENTS

```
RESULT 1
US-09-359-268A-3
; Sequence 3, Application US/09359268A
; Patent No. 6630294
; GENERAL INFORMATION:
; APPLICANT: Peltz, Stuart
; APPLICANT: Czaplinski, Kevin
; APPLICANT: Dimman, Jonathan D.
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
; FILE REFERENCE: 601-1-85N
; CURRENT APPLICATION NUMBER: US/09/359,268A
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,685
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; OTHER INFORMATION: Xaa = any amino acid
US-09-359-268A-3

Query Match      74.1% Score 40; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VVIDEXXQAXXXXXXIP1 17
DB      1 VVIDEXXQAXXXXXXIP1 17

RESULT 2
US-09-902-540-13564
; Sequence 13564, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: MYXOCOCCUS XANTHUS Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-110(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
```

NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 13564  
LENGTH: 513  
TYPE: PRT  
ORGANISM: Myxococcus xanthus  
US-09-902-540-13564

Query Match 57.4%; Score 31; DB 4; Length 513;  
Best Local Similarity 29.4%; Pred. No. 26;  
Matches 5; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 VVIDEXXQAXXXXXXIP1 17  
Db 53 VVDESSQCRLEBALPV 69

RESULT 3  
US-09-252-991A-27962  
Sequence 27962, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 27962  
LENGTH: 302  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27962

Query Match 55.6%; Score 30; DB 4; Length 302;  
Best Local Similarity 77.8%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQAXXXXXXIP1 17  
Db 70 VVIDELLOA 78

RESULT 4  
US-09-538-092-1100  
Sequence 1100, Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:  
APPLICANT: Glot, Loic  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: CuratSeqFormatter Version 0.9  
SEQ ID NO 1100  
LENGTH: 993  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (0)...(0)  
OTHER INFORMATION: Polypeptide Accession Number P38935  
US-09-538-092-1100

Query Match 55.6%; Score 30; DB 4; Length 993;  
Best Local Similarity 77.8%; Pred. No. 90;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQAXXXXXXIP1 17  
Db 372 VVIDECAQA 380

RESULT 5  
US-09-248-796A-20994  
Sequence 20994, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstock et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 20994  
LENGTH: 73  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-20994

Query Match 53.7%; Score 29; DB 4; Length 73;  
Best Local Similarity 37.5%; Pred. No. 10;  
Matches 6; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 VVIDEXXQAXXXXXXIP1 17  
Db 11 IISASRATSVPAIP1 26

RESULT 6  
US-09-252-991A-17847  
Sequence 17847, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 17847  
LENGTH: 161  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17847

Query Match 53.7%; Score 29; DB 4; Length 161;  
Best Local Similarity 31.2%; Pred. No. 24;  
Matches 5; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 VVIDEXXQAXXXXXXIP1 17  
Db 29 LVDEEPQALREBALPV 44

RESULT 7



```
US-09-359-268A-29
; Sequence 29, Application US/09359268A
; Patent No. 6630294
; GENERAL INFORMATION:
; APPLICANT: Peltz, Stuart
; APPLICANT: Czapinski, Kevin
; APPLICANT: Dliman, Jonathan D.
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 601-1-85N
; CURRENT APPLICATION NUMBER: US/09/359,268A
; CURRENT FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,685
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-359-268A-29

Query Match          53.7%; Score 29; DB 4; Length 380;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 VVIDEXXQAXXXXXXIP 16
DB      146 VLIDESTQA 154

RESULT 8
US-09-540-236-2529
; Sequence 2529, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NOCTIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAE
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2529
; LENGTH: 413
; TYPE: PRT
; ORGANISM: M.cattarrhalis
US-09-540-236-2529

Query Match          53.7%; Score 29; DB 4; Length 413;
Best Local Similarity 31.2%; Pred. No. 63;
Matches 5; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY      1 VVIDEXXQAXXXXXXIP 16
DB      211 IIFDEVLQAGDVLVYP 226

RESULT 9
US-09-902-540-11693
; Sequence 11693, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
```

```
US-09-902-540-11693
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 11693
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-11693

Query Match          53.7%; Score 29; DB 4; Length 671;
Best Local Similarity 33.3%; Pred. No. 1e+02;
Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      2 VVIDEXXQAXXXXXXIP 16
DB      655 VVIDELKASDAAYVP 669

RESULT 10
US-08-724-354D-22
; Sequence 22, Application US/08724354D
; Patent No. 5994119
; GENERAL INFORMATION:
; APPLICANT: Dietz, Harry C.
; TITLE OF INVENTION: MAMMALIAN REGULATOR OF
; TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,354D
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/016,482
; FILING DATE: 29-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hallie, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/090001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-678-5099
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 971 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-724-354D-22

Query Match          53.7%; Score 29; DB 2; Length 971;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 VVIDEXXQAXXXXXXIP 9
DB      569 VLIDESTQA 577

RESULT 11
US-09-270-984A-22
; Sequence 22, Application US/09270984A
; Patent No. 6048965
; GENERAL INFORMATION:
```

APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
NONSENSE-MEDIATED RNA DECAY  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/270,984A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/724,354  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 971 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-270-984A-22

Query Match 53.7%; Score 29; DB 3; Length 971;  
Best Local Similarity 66.7%; Pred. No. 1.5e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
Db 569 VLIDESTQA 577

RESULT 12  
US-09-177-431-8  
Sequence 8, Application US/09177431  
Patent No. 6071700  
GENERAL INFORMATION:  
APPLICANT: He, Feng  
APPLICANT: Jacobson, Allan S.  
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE  
TITLE OF INVENTION: ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/177,431  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/955,472

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Faese, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 07917/050001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-9806  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 971 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-09-177-431-8

Query Match 53.7%; Score 29; DB 3; Length 971;  
Best Local Similarity 66.7%; Pred. No. 1.5e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
Db 569 VLIDESTQA 577

RESULT 13  
US-09-248-796A-19347  
Sequence 9, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Ketch Weinstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 19347  
LENGTH: 917  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-19347

Query Match 52.8%; Score 28.5; DB 4; Length 917;  
Best Local Similarity 38.9%; Pred. No. 1.9e+02;  
Matches 7; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 1 VVIDEXXQAXX-XXXIP1 17  
Db 619 VIIDATQSSSEPTLIP1 636

RESULT 14  
US-08-724-354D-9  
Sequence 9, Application US/08724354D  
Patent No. 5994119  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
NONSENSE-MEDIATED RNA DECAY  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA

ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724.354D  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/016.482  
FILING DATE: 29-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38.347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-724-354D-9

Query Match 51.9%; Score 28; DB 2; Length 30;  
Best Local Similarity 55.6%; Pred. No. 7.1;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Cy 1 VVIDEXXQA 9  
Db 1 ILIDESTQA 9

RESULT 15  
US-09-270-984A-9  
Sequence 9, Application US/09270984A  
Patent No. 6048965  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/270.984A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/724.354  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38.347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-270-984A-9

Query Match 51.9%; Score 28; DB 3; Length 30;  
Best Local Similarity 55.6%; Pred. No. 7.1;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Cy 1 VVIDEXXQA 9  
Db 1 ILIDESTQA 9

Search completed: April 18, 2005, 08:18:33  
Job time : 22.219 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 18, 2005, 08:06:16 ; Search time 60.3066 Seconds  
(without alignments)  
93.693 Million cell updates/sec

Title: US-10-652-334-3

Perfect score: 54

Sequence: 1 VVIDEXXQAXXXXXXIP1 17

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

1421835

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications MA:\*

1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCTI\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	40	74.1	17	US-10-652-334-3
2	38	70.4	256	US-10-437-963-111797
3	38	70.4	1323	US-10-437-963-111794
4	37	68.5	163	US-10-767-701-99859
5	32.5	60.2	650	US-10-369-493-2884
6	32	59.3	332	US-10-425-114-38637
7	32	59.3	404	US-10-425-114-37833
8	32	59.3	710	US-10-282-122A-62169
9	32	59.3	1944	US-10-369-493-2521
10	32	59.3	1975	US-10-437-963-140079
11	31	57.4	1919	US-10-767-701-57601
12	31	57.4	262	US-10-424-599-252174
13	31	57.4	559	US-10-437-963-190072

14	31	57.4	626	US-10-437-963-166322
15	31	57.4	637	US-10-424-599-233501
16	31	57.4	639	US-10-425-114-37717
17	31	57.4	712	US-10-369-493-22762
18	31	57.4	830	US-10-369-493-4012
19	30	55.6	66	US-10-424-599-167034
20	30	55.6	282	US-10-156-761-10407
21	30	55.6	395	US-10-369-493-17471
22	30	55.6	439	US-10-767-701-41078
23	30	55.6	642	US-10-437-963-111267
24	30	55.6	642	US-10-425-114-37557
25	30	55.6	673	US-10-369-493-21435
26	30	55.6	673	US-10-424-599-195146
27	30	55.6	1027	US-10-437-963-185291
28	29	54.6	648	US-10-369-493-21357
29	29	53.7	199	US-10-767-701-38822
30	29	53.7	361	US-10-282-122A-47247
31	29	53.7	380	US-10-652-334-29
32	29	53.7	638	US-10-437-963-181690
33	29	53.7	734	US-10-424-599-205643
34	29	53.7	925	US-10-369-493-2087
35	28	51.9	134	US-10-424-599-186407
36	28	51.9	188	US-10-425-114-71289
37	28	51.9	260	US-10-369-493-17089
38	28	51.9	289	US-10-369-493-1216
39	28	51.9	339	US-10-282-122A-45661
40	28	51.9	340	US-10-282-122A-67195
41	28	51.9	342	US-10-282-122A-46769
42	28	51.9	377	US-09-910-430-27
43	28	51.9	377	US-10-165-605A-27
44	28	51.9	414	US-10-425-114-41672
45	28	51.9	510	US-10-437-963-173736

## ALIGNMENTS

RESULT 1  
US-10-652-334-3  
Sequence 3, Appl1  
Publication No. US20040115787A1  
GENERAL INFORMATION:  
APPLICANT: Pelcz, Stuart  
APPLICANT: Czaplinski, Kevin  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/10/652.334  
CURRENT FILING DATE: 2003-08-28  
PRIOR APPLICATION NUMBER: US/09/359,268A  
PRIOR FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
PRIOR FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
FEATURE:  
OTHER INFORMATION: Xaa = any amino acid  
US-10-652-334-3

Query Match 74.1%; Score 40; DB 16; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.0094;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VVIDEXXQAXXXXXXIP1 17  
Db 1 VVIDEXXQAXXXXXXIP1 17

## RESULT 2

US-10-437-963-111797  
; Sequence 111797, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 111797  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(256)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_15741C.1.pep  
US-10-437-963-111797

Query Match Best Local Similarity 70.4%; Score 38; DB 16; Length 256;  
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 VVIDEXXQAXXXXXXIP 16  
Db 137 VVIDEAAQASEVAVLP 152

## RESULT 3

US-10-437-963-111794  
; Sequence 111794, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 111794  
; LENGTH: 1323  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(1323)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_15739C.1.pep  
US-10-437-963-111794

Query Match Best Local Similarity 70.4%; Score 38; DB 16; Length 1323;

Best Local Similarity 50.0%; Pred. No. 4.2;  
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 VVIDEXXQAXXXXXXIP 16  
Db 993 VVIDEAAQASEVAVLP 1008

## RESULT 4

US-10-767-701-49859  
; Sequence 49859, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53353)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 49859  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3478-020-Pl-K1-F11.pep  
US-10-767-701-49859

Query Match Best Local Similarity 68.5%; Score 37; DB 16; Length 163;  
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 VVIDEXXQAXXXXXXIP 16  
Db 17 VVIDEAAQASEVAVLP 32

## RESULT 5

US-10-369-493-2884  
; Sequence 2884, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2884  
; LENGTH: 650  
; TYPE: PRT  
; ORGANISM: Thermotoga maritima  
US-10-369-493-2884

Query Match Best Local Similarity 60.2%; Score 32.5; DB 15; Length 650;  
Matches 9; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

Qy 1 VVIDEXXQAXX-XXIPI 17  
Db 380 VVIDEASQATIPSIILPI 397

## RESULT 6

```
US-10-425-114-38637
; Sequence 38637, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38637
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700241578_FLI.pep
US-10-425-114-38637

Query Match          59.3%; Score 32; DB 15; Length 332;
Best Local Similarity 77.8%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 VVIDEXXQA 9
      |||||
Db      58 VVIDEAAQA 66

RESULT 7
US-10-425-114-37833
; Sequence 37833, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37833
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700618821_FLI.pep
US-10-425-114-37833

Query Match          59.3%; Score 32; DB 15; Length 404;
Best Local Similarity 77.8%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 VVIDEXXQA 9
      |||||
Db      130 VVIDEAAQA 138

RESULT 8
US-10-282-122A-62169
; Sequence 62169, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liansu

; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62169
; LENGTH: 710
; TYPE: PRT
; ORGANISM: Mycobacterium avium
US-10-282-122A-62169

Query Match          59.3%; Score 32; DB 15; Length 710;
Best Local Similarity 31.2%; Pred. No. 62;
Matches 5; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy      2 VVIDEXXQXXXXXPI 17
      ::|||
Db      558 IMDETQAAADGJPL 573

RESULT 9
US-10-369-493-2521
; Sequence 2521, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2521
; LENGTH: 1944
```

TYPE: PRT  
ORGANISM: Schizosaccharomyces pombe  
US-10-369-493-2521

Query Match  
Best Local Similarity 46.7%; Score 32; DB 15; Length 1944;  
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 VVIDEXXQAXXXXI 15  
DB 1513 VVIDEAAQAVELSSI 1527

## RESULT 10

US-10-437-963-140079  
Sequence 140079, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 140079  
LENGTH: 1975  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_41310C.1.pdp  
US-10-437-963-140079

Query Match  
Best Local Similarity 59.3%; Score 32; DB 16; Length 1975;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
DB 1384 VVIDEAAQA 1392

## RESULT 11

US-10-767-701-57601  
Sequence 57601, Application US/10767701  
Publication No. US20040172684A1  
GENERAL INFORMATION:  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53235)B  
CURRENT APPLICATION NUMBER: US/10/767,701  
CURRENT FILING DATE: 2004-01-29  
NUMBER OF SEQ ID NOS: 63128  
SEQ ID NO 57601  
LENGTH: 219  
TYPE: PRT  
ORGANISM: Sorghum bicolor  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1) (219)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: 3096562.pdp

US-10-767-701-57601

Query Match  
Best Local Similarity 57.4%; Score 31; DB 16; Length 219;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
DB 50 VVIDEAAQA 58

## RESULT 12

US-10-424-599-252174  
Sequence 252174, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 252174  
LENGTH: 262  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1) (262)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_69741C.1.pdp  
US-10-424-599-252174

Query Match  
Best Local Similarity 57.4%; Score 31; DB 15; Length 262;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
DB 176 VVIDEAAQA 184

## RESULT 13

US-10-437-963-190072  
Sequence 190072, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 190072  
LENGTH: 559  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_8651C.1.pdp  
US-10-437-963-190072



Query Match 57.4%; Score 31; DB 16; Length 559;  
 Best Local Similarity 66.7%; Pred. No. 84;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
 ||:||||  
 DB 452 VVIDEAAQA 460

## RESULT 14

US-10-437-963-166322  
 ; Sequence 166322, Application US/10437963  
 ; Publication No. US20040123343A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LA ROSA, Thomas J.  
 ; APPLICANT: KOVALIC, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Mu, Wei  
 ; APPLICANT: Boucharov, Andrey A.  
 ; APPLICANT: Barbazuk, Brad  
 ; APPLICANT: Li, Ping  
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53221)B  
 ; CURRENT APPLICATION NUMBER: US/10/437,963  
 ; CURRENT FILING DATE: 2003-05-14  
 ; NUMBER OF SEQ ID NOS: 204966  
 ; SEQ ID NO 166322  
 ; LENGTH: 626  
 ; TYPE: PRT  
 ; ORGANISM: Oryza sativa  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(626)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_65042C.1.pep  
 ; US-10-437-963-166322

Query Match 57.4%; Score 31; DB 16; Length 626;  
 Best Local Similarity 66.7%; Pred. No. 95;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
 ||:||||  
 DB 353 VVIDEAAQA 361

## RESULT 15

US-10-424-599-233501  
 ; Sequence 233501, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LA ROSA, Thomas J.  
 ; APPLICANT: KOVALIC, David K.  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 233501  
 ; LENGTH: 637  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(637)  
 ; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_52878C.1.pep  
 ; US-10-424-599-233501

Query Match 57.4%; Score 31; DB 15; Length 637;  
 Best Local Similarity 66.7%; Pred. No. 97;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
 ||:||||  
 DB 363 VVIDEAAQA 371

Search completed: April 18, 2005, 09:04:06  
 Job time : 61.3066 secs

This Page Blank (uspro)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 18, 2005, 07:32:05 ; Search time 16.3796 Seconds  
(without alignments)  
99.861 Million cell updates/sec

Title: US-10-652-334-3  
Perfect score: 54  
Sequence: 1 VVIDEXXQAXXXXXXIP1 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	72.2	1113	2	E64215 hypothetical prote
2	39	72.2	1113	2	S73327 MG140 homolog - My
3	37	68.5	1090	2	T00533 probable DNA2-NM7
4	37	68.5	1311	2	T08986 hypothetical prote
5	32.5	60.2	650	2	G72429 hypothetical prote
6	32	59.3	1944	2	T40065 tRNA-splicing endo
7	32	59.3	2142	2	D86303 p17P16.1 protein -
8	31	57.4	530	2	D70476 DNA helicase - Agu
9	31	57.4	660	2	E71420 hypothetical prote
10	31	57.4	555	2	T41580 probable dna-bindi
11	31	57.4	1687	2	T39072 DNA2-NM7 helicase
12	30	55.6	265	2	B83395 probable enoyl-CoA
13	30	55.6	395	2	C64138 8-amino-7-oxononan
14	30	55.6	635	2	T02699 probable helicase
15	30	55.6	663	2	H63312 probable DNA helic
16	30	55.6	860	2	T15778 hypothetical prote
17	30	55.6	993	2	A47500 Ig mu chain switch
18	30	55.6	1060	1	A40264 kinesin-related pr
19	29.5	54.6	648	2	C68423 DNA helicase homol
20	29	53.7	156	2	E83227 hypothetical prote
21	29	53.7	191	2	T21031 hypothetical prote
22	29	53.7	203	2	T44695 butyr protein [impo
23	29	53.7	250	2	AB2404 hypothetical prote
24	29	53.7	361	2	B70189 rod shape-determin
25	29	53.7	496	2	S68160 probable RNA bindi
26	29	53.7	643	2	H84650 probable transcrip
27	29	53.7	935	2	S62476 hypothetical prote
28	29	53.7	971	2	S23408 prematurely termin
29	29	53.7	989	2	T48845 insulin II gene en

30	29	53.7	993	2	S35633 DNA-binding protei
31	29	53.7	1069	2	T43280 nonsense-mediated
32	28	51.9	121	2	AD2492 hypothetical prote
33	28	51.9	178	2	AC1815 cob (I) alamin adeno
34	28	51.9	260	2	G87712 hsp protein (cycl
35	28	51.9	270	2	C83829 phosphomethylpyrim
36	28	51.9	289	2	D69044 diaminopimelate ep
37	28	51.9	322	2	S75280 hypothetical prote
38	28	51.9	337	2	D69660 rod shape-determin
39	28	51.9	338	2	JC5436 aspartate-semialde
40	28	51.9	346	2	AC0302 lacI-family transac
41	28	51.9	468	2	E75313 conserved hypotnet
42	28	51.9	666	2	S01283 hypothetical prote
43	28	51.9	742	2	B86319 F25116.10 protein
44	28	51.9	745	2	G72453 hypothetical prote
45	28	51.9	924	2	A44945 104K microneme-zho

## ALIGNMENTS

## RESULT 1

E64215  
hypothetical protein MG140 - Mycoplasma genitalium  
C/Species: Mycoplasma genitalium  
C/Date: 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C/Accession: E64215, S186594  
R/Praser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;  
M.; Fuhmann, J.; Nguyen, D.; Ullrich, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.  
, C.A.; Venter, J.C.  
Science 270, 397-403, 1995  
A/Title: The minimal gene complement of Mycoplasma genitalium.  
A/Reference number: A64200; MUID:96026346; PMID:7569993  
A/Accession: E64215  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-1113 <TR>  
A/Cross-references: UNIPROT:P47386; GB:U39694; GB:U43967; NID:g1045822; PID:g1045823; TI  
A/Experimental source: strain G-37  
C/Genetics:  
A/Genetic code: SGC3

Query Match 72.2%, Score 39; DB 2; Length 1113;  
Best Local Similarity 52.9%, Pred. No. 0.35;  
Matches 9; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 VVIDEXXQAXXXXXXIP1 17  
Db 699 VVIDEXXQAXXXXXXIP1 715

## RESULT 2

S73327  
MG140 homolog - Mycoplasma pneumoniae (strain ATCC 29342)  
N/Alternate names: hypothetical protein E07\_orf113  
C/Species: Mycoplasma pneumoniae  
A/Variety: ATCC 29342  
C/Date: 26-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C/Accession: S73327  
R/Himmelreich, R.; Hiltbert, H.; Plagens, H.; Plick, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A/Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A/Reference number: S73327; MUID:97105885; PMID:8948633  
A/Accession: S73327  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-1113 <HM>  
A/Cross-references: UNIPROT:P75033; EMBL:AE000001; GB:U00089; NID:g1673645; PIDN:AA89564  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C/Genetics:  
A/Genetic code: SGC3

Query Match 72.2%, Score 39; DB 2; Length 1113;

Best Local Similarity 52.9%; Pred. No. 0.35;  
Matches 9; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 VVIDEXXQAXXXXXXIP 17  
Db 699 VVIDEASQVLEIRAIPI 715

## RESULT 3

T00533  
Probable DNA2-NAM7 helicase family protein [imported] - Arabidopsis thaliana

N;Alternate names: SEN1 protein homolog T20K24.14

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 09-Jul-2004

C;Accession: T00533; G84572

R;Rounbley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, July 1997

A;Description: Arabidopsis thaliana chromosome II BAC T20K24 genomic sequence.

A;Reference number: Z14167

A;Accession: T00533

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1090 <ROU>

A;Cross-references: UNIPROT:O64476; EMBL:AC002392; NID:G3176701; PID:G3176714

A;Experimental source: cultivar Columbia

R;Lin, X.; Kaul, S.; Rounbley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffett, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.B.; Unayan, L.; Tallon, L.; euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: G84572

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1090 <STO>

A;Cross-references: GB:AE002093; NID:G3176714; PID:ADJ12029.1; GSPDB:GN00139

C;Genetics:

A;Gene: T20K24.14; A2319120

A;Map position: 2

A;Introns: 250/1; 310/1; 342/1; 383/2; 427/3; 1016/3

Query Match 68.5%; Score 37; DB 2; Length 1090;

Best Local Similarity 50.0%; Pred. No. 1.1;

Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 VVIDEXXQAXXXXXXIP 16

Db 778 VVIDEASQVLEIRAIPI 793

RESULT 4

T08986  
Hypothetical protein F6G3.130 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C;Accession: T08986

R;Bayan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Meyer, K.F.X submitted to the Protein Sequence Database, May 1999

A;Reference number: Z16520

A;Accession: T08986

A;Molecule type: DNA

A;Residues: 1-1311 <BEV>

A;Cross-references: UNIPROT:Q9SCW3; EMBL:AL078464; GSPDB:GN00062; ATSP:F6G3.130

A;Experimental source: cultivar Columbia; BAC clone F6G3

C;Genetics:

A;Gene: ATSP:F6G3.130

A;Map position: 4

A;Introns: 414/1; 474/1; 506/1; 547/2; 591/3; 1179/3

Query Match 68.5%; Score 37; DB 2; Length 1311;

Best Local Similarity 50.0%; Pred. No. 1.3;

Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 VVIDEXXQAXXXXXXIP 16  
Db 941 VVIDEASQVLEIRAIPI 956

## RESULT 5

G72429  
Hypothetical protein TM0005 - Thermotoga maritima (strain MSB8)

C;Species: Thermotoga maritima

C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C;Accession: G72429

R;Neilson, K.E.; Clayton, R.A.; Gill, S.R.; Gwim, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A;Reference number: A72200; MUID:99287316; PMID:10360571

A;Accession: G72429

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-650 <ARN>

A;Cross-references: UNIPROT:Q9WXM0; GB:AE001689; GB:AE000512; NID:G4980483; PID:ADJ35095

A;Experimental source: strain MSB8

C;Genetics:

A;Gene: TM0005

C;Superfamily: probable DNA helicase MJ0104

Query Match 60.2%; Score 32.5; DB 2; Length 650;

Best Local Similarity 50.0%; Pred. No. 8.6;

Matches 9; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

Qy 1 VVIDEXXQAXX-XXXIP 17

Db 380 VVIDEASQATIPSLIPI 397

RESULT 6

T40065  
CRNA-splicing endonuclease positive effector - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C;Accession: T40065

R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; LeJaure, V.; Galibert, F.

A;Title: The EMBL Data Library, December 1998

A;Reference number: Z21903

A;Accession: T40065

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1944 <LYN>

A;Cross-references: UNIPROT:Q94387; EMBL:AL034463; PIDN:CAA22438.1; GSPDB:GN00067; SPDB:8

A;Experimental source: strain 972h; cosmid c29A10

C;Genetics:

A;Gene: SPDB:SPBC29A10.10C

A;Map position: 2

Query Match 59.3%; Score 32; DB 2; Length 1944;

Best Local Similarity 46.7%; Pred. No. 37;

Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 VVIDEXXQAXXXXXX 15

Db 1513 VVIDEASQVLEIRAIPI 1527

RESULT 7

D86303  
P17P16.1 protein - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: D86303

R;Theologis, A.; Becker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Hutzar, L.

Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L., Jenkins, J., Johnson-Hopson, C., Khan, S., Khaykin, E., Kim, C., C.A., Li, J.H., Li, X., Lin, K., Liu, S.X., Liu, Z.A., Luro, J.S., Malt, R., Mazzilli, Rizzo, M., Rooney, T., Rowley, D., Sakano, H.  
A:Authors: Salzberg, S.L., Schwartz, J.R., Shum, P., Southwick, A.M., Sun, H., Tallon, Ker, M., Wu, D., Yu, G., Fraser, C.M., Venter, J.C., Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: D86303  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2142 <STO>  
A:Cross-references: UNIPROT:Q9FWR3; GB:AB005172; NID:99954728; PID:AA09081.1; GSPDB:GN C:Genetics:  
A:Map position: 1

Query Match 59.3%; Score 31; DB 2; Length 2142;  
Best Local Similarity 77.8%; Pred. No. 41;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
Db 1526 VVIDEAAQA 1534

RESULT 8  
D70476  
DNA helicase - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C:Accession: D70476  
R:Decker, G., Warren, P.V., Gaasterland, T., Young, W.G., Lenox, A.L., Graham, D.E.; Oy V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: D70476  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-530 <ADP>  
A:Cross-references: UNIPROT:O67840; GB:AB000770; NID:92984274; PID:AA07803.1; PID:9298 A:Experimental source: strain VPS  
C:Genetics:  
A:Gene: helix

Query Match 57.4%; Score 31; DB 2; Length 530;  
Best Local Similarity 66.7%; Pred. No. 16;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
Db 259 VVIDEATQA 267

RESULT 9  
E71420  
hypotheetical protein - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
A:Variety: Columbia  
C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 09-Jul-2004  
C:Accession: E71420  
R:Bevan, M., Bancroft, I., Bent, E., Love, K., Goodman, H., Dean, C., Bergkamp, R., Dirk P., Wedler, H., Wedler, E., Wandut, R., Weitzenecker, T., Pohl, T.M., Terry, N., Giel avanagh, T., Hempel, S., Kotter, P., Entian, K.D., Rieger, M., Schaeffer, M., Funk, B.  
Nature 391, 485-488, 1998  
A:Authors: Mueller-Auer, S., Silvey, M., James, R., Montfort, A., Pons, A., Puldomenect erhoff, A., Moore, T., Jones, J.D.G., Eneva, T., Palme, K., Bene, V., Rechman, S., Ans C., Chavakis, N.  
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal A:Reference number: A71400; MUID:98121113; PMID:9461215  
A:Accession: E71420  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A:Residues: 1-555 <BR>  
A:Cross-references: UNIPROT:O23408; GB:Z97339; NID:92244901; PID:e326936; PID:92244913 C:Genetics:  
A:Map position: 4COP-463845

Query Match 57.4%; Score 31; DB 2; Length 555;  
Best Local Similarity 66.7%; Pred. No. 17;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
Db 229 VVIDEAAQA 237

RESULT 10  
T41580  
Probable dna-binding protein - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T41580  
R:Murphy, L., Harris, D., Wood, V., Lyne, M.H., Rajandream, M.A., Barrell, B.G.  
Submitted to the EMBL Data Library, September 1998  
A:Reference number: Z22002  
A:Accession: T41580  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-660 <MR>  
A:Cross-references: UNIPROT:O94247; EMBL:AL011546; PID:CAA20863.1; GSPDB:GN00068; SPDB:S A:Experimental source: strain 972h(-)  
C:Genetics:  
A:Map position: 3  
A:Gene: SPDB:SPC737.07C  
A:Introns: 14/3; 72/2; 151/2  
C:Superfamily: probable DNA helicase MJ0104

Query Match 57.4%; Score 31; DB 2; Length 660;  
Best Local Similarity 66.7%; Pred. No. 21;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
Db 386 VVIDEASQA 394

RESULT 11  
T39072  
DNA2-NAM7 helicase family protein - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T39072  
R:Murphy, L., Harris, D., Barrell, B.G., Rajandream, M.A., Connor, R.E., Wood, V.  
Submitted to the EMBL Data Library, August 1997  
A:Reference number: Z21825  
A:Accession: T39072  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1687 <MR>  
A:Cross-references: UNIPROT:Q92355; EMBL:Z81317; PID:CA03612.1; GSPDB:GN00066; SPDB:SP A:Experimental source: strain 972h-, cosmid c6G9  
C:Genetics:  
A:Gene: SPDB:SPAC69.010C  
A:Map position: 1

Query Match 57.4%; Score 31; DB 2; Length 1687;  
Best Local Similarity 66.7%; Pred. No. 57;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVIDEXXQA 9  
Db 1375 VVIDEAAQA 1383

RESULT 12

B83395  
 Probable enoyl-CoA hydratase/isomerase PA2013 [imported] - *Pseudomonas aeruginosa* (strain C) [Species: *Pseudomonas aeruginosa*]  
 C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 16-Aug-2004  
 C/Accession: B83395  
 R/Source: C.K.; Pam, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B. adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim, J.; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A/Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
 A/Reference number: A82950; MUID:20437337; PMID:10984043  
 A/Accession: B83395  
 A/Status: Preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-265 <STO>  
 A/Cross-references: UNIPROT:O91298; GB:AE004627; GB:AE004091; NID:g9948007; PIDD:AA0540  
 A/Experimental source: strain PA01  
 C/Genetics:  
 A/Gene: PA2013  
 C/Superfamily: Naphthoate synthase; enoyl-CoA hydratase homology

Query Match 55.6%; Score 30; DB 2; Length 265;  
 Best Local Similarity 77.8%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 VVIDEXXQA 9  
 DB 33 VVIDELQA 41

RESULT 13  
 C84138  
 8-amin-7-oxononanoate synthase biof [imported] - *Bacillus halodurans* (strain C-125)  
 C/Species: *Bacillus halodurans*  
 C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
 C/Accession: C84138  
 R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
 A/Reference number: A83650; MUID:20512582; PMID:11058132  
 A/Accession: C84138  
 A/Status: Preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-395 <STO>  
 A/Cross-references: UNIPROT:O9K625; GB:AP001520; GB:BA000004; NID:g10176401; PIDD:BA076  
 A/Experimental source: strain C-125  
 C/Genetics:  
 A/Gene: biof  
 C/Superfamily: 5-aminolevulinatase synthase; glycine C-acetyltransferase homology  
 C/Keywords: phosphoprotein; pyridoxal phosphate (lys) (covalent) #status predicted  
 F/240/Binding site: pyridoxal phosphate (lys)

Query Match 55.6%; Score 30; DB 2; Length 395;  
 Best Local Similarity 29.4%; Pred. No. 21;  
 Matches 5; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 1 VVIDEXXQAXXXXP 17  
 DB 334 LVVBSLQAGIAAPV 350

RESULT 14  
 T02699  
 Probable helicase At2g03270 [imported] - *Arabidopsis thaliana*  
 N/Alternate names: hypothetical protein T18E12.6  
 C/Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C/Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
 C/Accession: T02699; D84446  
 R/Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Ro  
 submitted to the EMBL Data Library, September 1998  
 A/Description: *Arabidopsis thaliana* chromosome II BAC T18E12 genomic sequence.  
 A/Reference number: Z14702  
 A/Accession: T02699

A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-635 <RNA>  
 A/Cross-references: UNIPROT:O81047; EMBL:AC005313; NID:g3548797; PID:g3548803  
 A/Experimental source: cultivar Columbia  
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffett, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Unayama, L.; Tallon, L.;  
 euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
 Nature 402, 761-768, 1999  
 A/Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
 A/Reference number: A84420; MUID:20083487; PMID:10617197  
 A/Accession: D84446  
 A/Status: Preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-635 <STO>  
 A/Cross-references: GB:AE002093; NID:g4335770; PIDD:AA017447.1; GSPDB:GN00139  
 C/Genetics:  
 A/Gene: At2g03270; T18E12.6  
 A/Map position: 2  
 C/Superfamily: probable DNA helicase MJ0104

Query Match 55.6%; Score 30; DB 2; Length 635;  
 Best Local Similarity 66.7%; Pred. No. 36;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VVIDEXXQA 9  
 DB 362 VVIDEGQA 370

RESULT 15  
 H64312  
 Probable DNA helicase MJ0104 - *Methanococcus jannaschii*  
 C/Species: *Methanococcus jannaschii*  
 C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
 C/Accession: H64312  
 R/Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
 Reich, C.I.; Overbeek, R.; Kirsch, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
 reon, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
 A/Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*  
 A/Reference number: A64300; MUID:96337999; PMID:8688087  
 A/Accession: H64312  
 A/Status: Preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-663 <BLU>  
 A/Cross-references: UNIPROT:Q57568; GB:U67467; GB:L77117; NID:g2826242; PIDD:AA08084.1;  
 C/Genetics:  
 A/Map position: FOR99243-101234  
 A/Start codon: TTG  
 C/Superfamily: probable DNA helicase MJ0104

Query Match 55.6%; Score 30; DB 2; Length 663;  
 Best Local Similarity 66.7%; Pred. No. 37;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VVIDEXXQA 9  
 DB 398 VVIDEGQA 406

Search completed: April 18, 2005, 08:06:00  
 Job time : 17.3796 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 07:33:55 ; Search time 70.9781 Seconds  
(without alignments)  
122.648 Million cell updates/sec

Title: US-10-652-334-3  
Perfect score: 54  
Sequence: 1 VVIDEXXQAXXXXXXIRPI 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	72.2	1113	Y140_MYCG	P47386 mycoplasma
2	39	72.2	1113	Y140_MYCPN	P75033 mycoplasma
3	38	70.4	1468	Q8W5K9	Q8W5K9 oryza sativ
4	38	70.4	1468	Q7XHS9	Q7XHS9 oryza sativ
5	37	68.5	1090	Q64476	Q64476 arabidopsis
6	37	68.5	1311	Q9SZW3	Q9SZW3 arabidopsis
7	36	66.7	907	Q8EAB9	Q8EAB9 shewanella
8	33.5	62.0	967	Q8EAS0	Q8EAS0 dictyostell
9	33	61.1	1132	Q7KCL4	Q7KCL4 mycoplasma
10	33	61.1	1182	Q6Q9H5	Q6Q9H5 mycoplasma
11	32.5	60.2	650	Q9WXM0	Q9WXM0 thermotoga
12	32	59.3	318	Q8DWP8	Q8DWP8 streptococc
13	32	59.3	318	Q8E2J6	Q8E2J6 streptococc
14	32	59.3	405	Q6A0T5	Q6A0T5 desulfotale
15	32	59.3	1944	Q94387	Q94387 schizosacch
16	32	59.3	2142	Q9FWK1	Q9FWK1 arabidopsis
17	31	57.4	365	Q7KMW1	Q7KMW1 dictyostell
18	31	57.4	530	Q67840	Q67840 aquifex ae
19	31	57.4	555	Q23408	Q23408 arabidopsis
20	31	57.4	651	Q6E2U8	Q6E2U8 oryza sativ
21	31	57.4	660	Q94247	Q94247 schizosacch
22	31	57.4	716	Q7SIL9	Q7SIL9 neurospora
23	31	57.4	770	Q7X684	Q7X684 oryza sativ
24	31	57.4	788	Q851T6	Q851T6 oryza sativ
25	31	57.4	1687	1 SEW1_SCHPO	Q93355 schizosacch
26	30.5	56.5	1024	Q75JJO	Q75JJO dictyostell
27	30	55.6	145	Q8EH34	Q8EH34 shewanella
28	30	55.6	176	Q74C04	Q74C04 geobacter s
29	30	55.6	265	Q91298	Q91298 pseudomonas
30	30	55.6	282	Q8J395	Q8J395 streptomyce
31	30	55.6	395	Q9K625	Q9K625 bacillus ha

32	30	55.6	402	2	Q7VKG3	Q7VKG3 haemophilus
33	30	55.6	421	2	Q62MQ2	Q62MQ2 burkholderi
34	30	55.6	426	2	Q63XA8	Q63XA8 burkholderi
35	30	55.6	455	2	Q6CV89	Q6CV89 kluyveromyce
36	30	55.6	503	2	Q9RK46	Q9RK46 streptomyce
37	30	55.6	635	2	Q9FNX9	Q9FNX9 arabidopsis
38	30	55.6	639	2	Q81047	Q81047 arabidopsis
39	30	55.6	648	2	Q18479	Q18479 caenorhabdi
40	30	55.6	663	1	Y104_MERVA	Q57568 methanococc
41	30	55.6	794	2	Q6VYI6	Q6VYI6 oryza sativ
42	30	55.6	976	2	Q7RBP5	Q7RBP5 plasmodium
43	30	55.6	993	1	SMB2_HUMAN	P38935 homo sapien
44	30	55.6	1034	2	Q7N378	Q7N378 photorhabdu
45	30	55.6	1040	2	Q7P6Q0	Q7P6Q0 fusobacteri

## ALIGNMENTS

RESULT 1  
Y140\_MYCG STANDARD: PRT: 1113 AA.  
ID P47386; Q49203; Q49283; Q49445; Q49472;  
AC P47386; Q49203; Q49283; Q49445; Q49472;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Hypothetical ATP-binding protein MG140.  
GN OrderedLocustName=MG140;  
OS Mycoplasma genitalium.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 33530 / G-37;  
RX MEDLINE=96026346; PubMed=7569993;  
RA Frieser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G.G., Kelley J.M.,  
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,  
RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
RA Tomb J.-P., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,  
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
RT "The minimal gene complement of Mycoplasma genitalium.";  
RL Science 270:397-403 (1995).  
[2]  
RN SEQUENCE OF 206-267 FROM N.A.  
RP STRAIN=ATCC 33530 / G-37;  
RX MEDLINE=92051396; PubMed=1945886;  
RA Peterson S.N., Schramm N., Hu P.-C., Bort K.F., Hutchison C.A. III;  
RT "A random sequencing approach for placing markers on the physical map  
of Mycoplasma genitalium.";  
RL Nucleic Acids Res. 19:6027-6031 (1991).  
[3]  
RN SEQUENCE OF 128-267, 294-394 AND 999-1100 FROM N.A.  
RP STRAIN=ATCC 33530 / G-37;  
RX MEDLINE=94075230; PubMed=8253680;  
RA Peterson S.N., Hu P.-C., Bort K.F., Hutchison C.A. III;  
RT "A survey of the Mycoplasma genitalium genome by using random  
sequencing.";  
RL J. Bacteriol. 175:7918-7930 (1993).  
[4]  
RN -1-SIMILARITY: Belong to the DNA2/NAM7 helicase family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: U39695; AAC71358.1; -  
DR EMBL: X61512; CA43726.1; -  
DR EMBL: U01729; AAC4307.1; -  
DR EMBL: U01742; AAD10552.1; -

DR EMBL: U02156; AAD12438..1; -.  
 DR PIR: E64215; E64215.  
 DR TIGR: MG140; -.  
 KW ATP-binding, Complete proteome; Helicase; Hypothetical protein.  
 FT NP BIND 313 320 ATP (Potential).  
 FT CONFLICT 267 267 D -> R (in Ref. 2 and 3).  
 FT CONFLICT 394 394 L -> W (in Ref. 3).  
 FT CONFLICT 1092 1100 PIGVSKIR -> OLGWFLKSD (in Ref. 3).  
 SQ SEQUENCE 1113 AA; 130579 MW; FFOC51F926T253F CRC64;

Query Match 72.2%; Score 39; DB 1; Length 1113;  
 Best Local Similarity 52.9%; Pred. No. 2.5;  
 Matches 9; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 VVIDEXXQAXXXXXXIP 17  
 |||||  
 DB 699 VVIDEASQVLEIRAPI 715

RESULT 2  
 Y140\_MYCPN STANDARD; PRT; 1113 AA.  
 ID Y140\_MYCPN  
 AC P75033;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE Hypothetical ATP-binding protein MG140 homolog (E07\_orf1113).  
 GN OrderedLocustNames=MFN153; ORFNames=MF001;  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxId=2104;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC MEDLINE=97105885; PubMed=8948633; DOI=10.1093/nar/24.22.4420;  
 RA Himmelfreid R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,  
 RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae";  
 RL Nucleic Acids Res. 24:4420-4449 (1996).  
 CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AE000001; AAB95649.1; -.  
 DR PIR: S73327; S73327.  
 KW ATP-binding, Complete proteome; Helicase; Hypothetical protein.  
 FT NP BIND 313 320 ATP (Potential).  
 SQ SEQUENCE 1113 AA; 130333 MW; 48A3337EB081A40 CRC64;

Query Match 72.2%; Score 39; DB 1; Length 1113;  
 Best Local Similarity 52.9%; Pred. No. 2.5;  
 Matches 9; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 VVIDEXXQAXXXXXXIP 17  
 |||||  
 DB 699 VVIDEASQVLEIRAPI 715

RESULT 3  
 Q8W5K9 PRELIMINARY; PRT; 1468 AA.  
 AC Q8W5K9;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Putative DNA2-NAM7 helicase family protein.

GN Name=OSJNBa0079B05.7;  
 OS Oryza sativa (Rice).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxId=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wang R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,  
 RA Saeki C., Henry D., Oates R., Simmons J.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC079179; ALJ1152.1; -.  
 DR Gramene; Q8W5K9; -.  
 DR GO: GO:0004386; F:helicase activity; IEA.

Query Match 70.4%; Score 38; DB 2; Length 1468;  
 Best Local Similarity 50.0%; Pred. No. 5.7;  
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 VVIDEXXQAXXXXXXIP 16  
 |||||  
 DB 993 VVIDEAAQASEVAVLP 1008

RESULT 4  
 Q7XH59 PRELIMINARY; PRT; 1468 AA.  
 ID Q7XH59  
 AC Q7XH59;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Putative DNA2-NAM7 helicase family protein.  
 GN ORFNames=OSJNBa0079B05.7;  
 OS Oryza sativa (Japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxId=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA The Rice Chromosome 10 Sequencing Consortium;  
 RT "In-depth view of structure, activity, and evolution of rice  
 RT chromosome 10.";  
 RL Science 300:1566-1569 (2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Buehl C.R., Wang R.A., McComble W.R., Messing J., Yuan Q.;  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE017050; AAP51895.1; -.  
 DR Gramene; Q7XH59; -.  
 DR GO: GO:0004386; F:helicase activity; IEA.  
 KW Helicase.  
 SQ SEQUENCE 1468 AA; 163521 MW; B611B24B42A302D8 CRC64;

Query Match 70.4%; Score 38; DB 2; Length 1468;  
 Best Local Similarity 50.0%; Pred. No. 5.7;  
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 VVIDEXXQAXXXXXXIP 16  
 |||||  
 DB 993 VVIDEAAQASEVAVLP 1008

RESULT 5  
 O64476 PRELIMINARY; PRT; 1090 AA.  
 AC O64476;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Putative DNA2-NAM7 helicase family protein.



```

GN Name=At2g19120;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
RA Brandon R.C., Sykes S.M., Mason T.M., Kervlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002392; RAD12029.1; -.
DR PIR; T00533; T00533.
DR GO; GO:0004386; F:helicase activity; IEA.
KW Helicase.
SQ SEQUENCE 1090 AA; 121521 MW; EB170342E18DCA1P CRC64;

Query Match 68.5%; Score 37; DB 2; Length 1090;
Best Local Similarity 50.0%; Pred. No. 7.5;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Cy 1 VVIDEXXQAXXXXXXIP 16
Db 778 VVIDEAAQASVGVLP 793

RESULT 6
Q9SZW3 PRELIMINARY; PRT; 1311 AA.
AC Q9SZW3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 05-JUN-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical protein F6G3.130 (Hypothetical protein Ar4g30100).
GN Name=F6G3.130; Synonym=ATG30100;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Beyer M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL078464; CAB3845.1; -.
DR EMBL; AL161576; CAB1003.1; -.
DR PIR; T08986; T08986.
KW Hypothetical protein.
SQ SEQUENCE 1311 AA; 145624 MW; 0F430BEB3A02AA4 CRC64;

Query Match 68.5%; Score 37; DB 2; Length 1311;
Best Local Similarity 50.0%; Pred. No. 9.1;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Cy 1 VVIDEXXQAXXXXXXIP 16
Db 1 VVIDEXXQAXXXXXXIP 16

```

```

Db 941 VVIDEAAQASVGVLP 956

RESULT 7
Q8EA89 PRELIMINARY; PRT; 907 AA.
AC Q8EA89;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein SO4019;
GN OrderedLocustNames=SO4019;
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Shewanellaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=2229786; PubMed=12368813; DOI=10.1038/nbt749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.B., Galdos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N.L., Meche B.A.,
RA Clayton R.A., Meyer T., Tsaplin A., Scott J., Beanan M.J.,
RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Haft D.H., Kolonay J.F., Madupu R., Peterson J.D., Umayam L.A.,
RA White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Impraim M.,
RA Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA Venter J.C., Nealsen K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015833; AANS6993.1; -.
DR TIGR; SO4019; -.
DR InterPro; IPR001261; ARGE_dapE.
DR PROSITE; PS00758; ARGE_DAPB_CPG2.1; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 907 AA; 101616 MW; BB2DDDD9EB55307 CRC64;

Query Match 66.7%; Score 36; DB 2; Length 907;
Best Local Similarity 43.8%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Cy 1 VVIDEXXQAXXXXXXIP 16
Db 527 VVIDEATQANIACLP 542

RESULT 8
Q8GAS0 PRELIMINARY; PRT; 967 AA.
AC Q8GAS0;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Similar to Neurospora crassa. Related to SNI1 protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Richinger L., Szatranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum."
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;

```

RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC116956; AA051157.1; -

DR GO: GO:0005524; F:ATP binding; IEA.

DR GO: GO:0000166; F:nucleotide binding; IEA.

DR InterPro: IPR003593; AAA\_ATPase.

DR SMART: SM00382; AAA.1.

KM ATP-binding.

SO SEQUENCE 967 AA; 109236 MW; DD1CA69BDE696AA2 CRC64;

Query Match 62.0%; Score 33.5; DB 2; Length 967;

Best Local Similarity 50.0%; Pred. No. 49;

Matches 9; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

Qy 1 VVIDEXXQAXXXXXXIP 17

Db 588 VVIDEAAQAVETSTLIP 605

# RESULT 9

ID Q7NC14 PRELIMINARY; PRT; 1132 AA.

AC Q7NC14; 01-MAR-2004 (TREMBlrel. 26; Created)

DT 01-MAR-2004 (TREMBlrel. 26; Last sequence update)

DE 01-MAR-2004 (TREMBlrel. 26; Last annotation update)

DE UvrD.

GN ORFNames=MGA\_0793;

OS Mycoplasma gallisepticum.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI\_TaxID=2096;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=R.

RX MEDLINE=22830409; PubMed=12949158; DOI=10.1099/mic.0.26427-0;

RA Papazisi L., Gorton T.S., Kutish G., Matham P.F., Browning G.F.,

RA Nguyen D.K., Swartzell S., Madan A., Mahairas G., Geary S.J.;

RT "The complete genome sequence of the avian pathogen Mycoplasma

RT gallisepticum strain R(low)."

RT Microbiology 149:2307-2316(2003).

RL EMBL: AE016967; AAF56445.1; -

DR GO: GO:0005524; F:ATP binding; IEA.

DR GO: GO:0004003; F:ATP-dependent DNA helicase activity; IEA.

DR GO: GO:0003677; F:DNA binding; IEA.

DR GO: GO:0006281; P:DNA repair; IEA.

DR InterPro: IPR000212; UvrD-helicase.

DR Pfam: PF00580; UvrD-helicase; 1.

KM Complete proteome.

SO SEQUENCE 1132 AA; 132955 MW; B96BB7D52BDC211E CRC64;

Query Match 61.1%; Score 33; DB 2; Length 1132;

Best Local Similarity 37.5%; Pred. No. 77;

Matches 6; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 VVIDEXXQAXXXXXXIP 16

Db 713 VVIDEASQVFLERAI 728

# RESULT 10

Q6Q9H5

ID Q6Q9H5 PRELIMINARY; PRT; 1182 AA.

AC Q6Q9H5; 05-JUL-2004 (TREMBlrel. 27; Created)

DT 05-JUL-2004 (TREMBlrel. 27; Last sequence update)

DE 05-JUL-2004 (TREMBlrel. 27; Last annotation update)

DE ATP binding protein.

OS Mycoplasma hyopneumoniae.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI\_TaxID=2099;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=F7;

RA Stakenborg T., Fil A., Butaye P.;

RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY551928; AAS58472.1; -

SO SEQUENCE 1182 AA; 137806 MW; FB752EDCFEF78F8 CRC64;

Query Match 61.1%; Score 33; DB 2; Length 1182;

Best Local Similarity 35.3%; Pred. No. 80;

Matches 6; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 VVIDEXXQAXXXXXXIP 17

Db 667 VVIDEASQVFLERAI 683

# RESULT 11

ID Q9WXM0 PRELIMINARY; PRT; 650 AA.

AC Q9WXM0; 01-NOV-1999 (TREMBlrel. 12; Created)

DT 01-NOV-1999 (TREMBlrel. 12; Last sequence update)

DE 01-MAR-2004 (TREMBlrel. 26; Last annotation update)

DE DNA helicase, putative.

GN OrderedLocustNames=TM0005;

OS Thermotoga maritima.

OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.

OX NCBI\_TaxID=2336;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;

RX MEDLINE=99287316; PubMed=10360571; DOI=10.1038/20601;

RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,

RA Haft D.H., Hickey B.K., Peterson J.D., Nelson W.C., Ketchum K.A.,

RA McDonald L.A., Uitterback T.R., Malek J.A., Linher K.D., Garrett M.M.,

RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.L.,

RA Heidelberg J.F., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,

RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;

RT "Evidence for lateral gene transfer between Archaea and Bacteria from

RT genome sequence of Thermotoga maritima."

RL Nature 399:323-329(1999).

RL EMBL: AE001689; AAD35099.1; -

DR PTR: G72429; G72429.

DR TIGR: TM0005; -

DR GO: GO:0005524; F:ATP binding; IEA.

DR GO: GO:0003677; F:DNA binding; IEA.

DR GO: GO:0004386; F:helicase activity; IEA.

DR InterPro: IPR004483; put\_DNA\_helic.

DR TIGRFAMs: TIGR00376; put\_DNA\_helic; 1.

KM Complete proteome; Helicase.

SO SEQUENCE 650 AA; 74241 MW; 9378FED72603B2DF CRC64;

Query Match 60.2%; Score 32.5; DB 2; Length 650;

Best Local Similarity 50.0%; Pred. No. 59;

Matches 9; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

Qy 1 VVIDEXXQAXX-XXXIP 17

Db 380 VVIDEASQATIPILIP 397

# RESULT 12

Q8DWP8

ID Q8DWP8 PRELIMINARY; PRT; 318 AA.

AC Q8DWP8; 01-MAR-2003 (TREMBlrel. 23; Created)

DT 01-MAR-2003 (TREMBlrel. 23; Last sequence update)

DE 01-MAR-2004 (TREMBlrel. 26; Last annotation update)

DE Cardamate kinase.

GN Name=arcC-2; OrderedLocustNames=SA02167;

OS Streptococcus agalactiae (serotype V).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI\_TaxID=216466;

RN [1]

RP SEQUENCE FROM N.A.

```

RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547; DOI=10.1073/pnas.182380799;
RA Tettehl H., Maignant V., Cieslewicz M.J., Eiken J.A., Peterson S.N.,
RA Weseloh M.R., Paulsen I.T., Nelson K.B., Margatit I., Reed T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,
RA Carly H.A., Cline R.T., Van Aken S.B., Gill J., Scarselli M., Mora M.,
RA Jacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaldo D., Kappoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR EMBL, AE014289; JN01025.1; -.
DR HSSP, P35836; 1B7B.
DR TIGR, SAG2167; -.
DR GO, GO:0008804; F:carbamate kinase activity; IEA.
DR GO, GO:0008652; P:amino acid biosynthesis; IEA.
DR GO, GO:0006525; P:arginine metabolism; IEA.
DR InterPro, IPR01048; Aa_kinase.
DR InterPro, IPR003964; Bac_carb_kinase.
DR Pfam, PF00696; Aa_kinase; 1.
DR PRINTS, PRSFP00723; Carbamate_kin; 1.
DR PRINTS, PR01469; CARBMTKINASE.
DR TIGRFAMs, TIGR00746; arcc; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 318 AA; 34006 MW; 563471C8337F8C7 CRC64;

Query Match 59.3%; Score 32; DB 2; Length 318;
Best Local Similarity 41.2%; Pred. No. 38;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 VVIDEXQAXXXXXXPI 17
Db 117 VVIDEKQQAFTNPTKPI 133

RESULT 13
Q8E2J6 PRELIMINARY; PRT; 318 AA.
AC Q8E2J6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein gbs2126;
GN OrderedLocusNames=gbs2126;
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus;
OC NCBI_TaxID=216495;
CX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusnik C., Buchrieser C., Chevalier F., Frangoul L.,
RA Masdek T., Zouine M., Couve E., Laloui L., Poyart C., Tien-Cuc P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513(2002).
DR EMBL, AL766856; CAD47785.1; -.
DR HSSP, P35836; 1B7B.
DR SAGALLst; gbs2126; -.
DR GO, GO:0008804; F:carbamate kinase activity; IEA.
DR GO, GO:0008652; P:amino acid biosynthesis; IEA.
DR GO, GO:0006525; P:arginine metabolism; IEA.
DR InterPro, IPR01048; Aa_kinase.
DR InterPro, IPR003964; Bac_carb_kinase.
DR Pfam, PF00696; Aa_kinase; 1.
DR PRINTS, PR01469; CARBMTKINASE.
DR TIGRFAMs, TIGR00746; arcc; 1.
KW Complete proteome; Transferase.

```

```

SQ SEQUENCE 318 AA; 34006 MW; 563471C8337F8C7 CRC64;

Query Match 59.3%; Score 32; DB 2; Length 318;
Best Local Similarity 41.2%; Pred. No. 38;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 VVIDEXQAXXXXXXPI 17
Db 117 VVIDEKQQAFTNPTKPI 133

RESULT 14
Q6AQT5 PRELIMINARY; PRT; 405 AA.
ID Q6AQT5;
AC Q6AQT5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Probable acetate kinase.
GN OrderedLocusNames=DP0559;
OS Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
OC Desulfobulbaceae; Desulfocales.
OX NCBI_TaxID=84980;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LSV54 / DSM 12343;
RX PubMed=1530514;
RA Radus R., Ruepp A., Frickey T., Ratzel T., Partmann B., Stark M.,
RA Bauer M., Zibat A., Lombardot T., Becker I., Aumann J., Gellner K.,
RA Teeling H., Leuschner W.D., Gloeckner P.-O., Lupas A.N., Aumann R.,
RA Kleink H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT from permanently cold Arctic sediments.";
RL Environ. Microbiol. 6:887-902(2004).
CC -1- SIMILARITY: Belongs to the acetate kinase family.
DR EMBL, CR522870; CAG35288.1; -.
DR GO, GO:0006522; C:intracellular; IEA.
DR GO, GO:0016301; F:kinase activity; IEA.
DR GO, GO:0016774; F:phosphotransferase activity, carboxyl group. .; IEA.
DR GO, GO:0006082; P:organic acid metabolism; IEA.
DR GO, GO:0016310; P:phosphorylation; IEA.
DR InterPro, IPR000890; Acetate_kin.
DR InterPro, IPR004372; Acka.
DR Pfam, PF00871; Acetate_kinase; 1.
DR PRINTS, PR00471; ACETATEKINASE.
DR TIGRFAMs, TIGR00016; ackA; 1.
DR PROSITE, PS01075; ACETATE_KINASE_1; 1.
DR PROSITE, PS01076; ACETATE_KINASE_2; 1.
KW Complete proteome; Kinase; Transferase.
SQ SEQUENCE 405 AA; 44128 MW; ED790BC85559D715 CRC64;

Query Match 59.3%; Score 32; DB 2; Length 405;
Best Local Similarity 41.2%; Pred. No. 49;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 VVIDEXQAXXXXXXPI 17
Db 110 VVIDEKVIAITAIPL 126

RESULT 15
Q94387 PRELIMINARY; PRT; 1944 AA.
ID Q94387;
AC Q94387;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SPEC2A10.10c protein.
GN Name=SPBC29A10.10c;
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;

```

OC Schizosaccharomyces.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972h-;  
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown K., Brown S., Chillingworth T., Church C.,  
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
RA James K., Jones L., Jones M., Leacher S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,  
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinovitch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,  
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Muller-Auer S.,  
RA Gabel C., Fuchs M., Dusterhoft A., Filtz C., Holzer E., Moestl D.,  
RA Hilbert H., Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R.,  
RA Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Punnelle B.,  
RA Goffeau A., Cadieu S., Xiang Z., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Rhode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Useery D., Barrett B.G., Nurse P.,  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
DR EMBL: AL034463; CAA22438.1; -.  
DR PIR: T40065; T40065.  
DR GeneDB\_Spomb; SPBC29A10.10c; -.  
DR InterPro; IPR000873; AMP-bind.  
DR PROSITE; PS00455; AMP BINDING.  
SQ SEQUENCE 1944 AA; 222209 MW; 12B005A934BEE11C CRC64;

Query Match 59.3%; Score 32; DB 2; Length 1944;  
Best Local Similarity 46.7%; Pred. No. 2.3e+02;  
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 1 VVIDEXXOAXXXXXX1 15  
|:|:|:|:|:|:|  
Db 1513 VVIDEAAQAVELSSI 1527

Search completed: April 18, 2005, 08:15:41  
Job time : 72.9781 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: April 18, 2005, 07:10:15 ; Search time 59.7372 Seconds  
(without alignments)  
77.692 Million cell updates/sec

Title: US-10-652-334-4  
Perfect score: 44  
Sequence: 1 XXILAGDXQLP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	86.4	12	7	ABW01180
2	38	86.4	12	8	ADP44106 Saccharom
3	38	86.4	13	3	AY77807 Molif IV
4	38	86.4	611	8	AD521255 Bacteri
5	38	86.4	642	8	AD543096 Bacteri
6	38	86.4	648	8	AD542927 Bacteri
7	38	86.4	648	8	AD541904 Bacteri
8	38	86.4	655	8	AB62031 Recombina
9	38	86.4	656	8	ADN18615 Bacteri
10	38	86.4	988	6	AA026745 988-mex r
11	37	84.1	414	3	AA77816 Yeast Dip
12	37	84.1	414	7	ABW01204 Saccharom
13	37	84.1	414	8	ADP44130 Yeast hel
14	37	84.1	650	8	ADN20231 Bacteri
15	37	84.1	653	8	AD543215 Bacteri
16	37	84.1	654	8	AA896566 Putative
17	37	84.1	683	8	AD543834 Bacteri
18	37	84.1	688	8	ADN47941 Thermoco
19	37	84.1	830	8	ADN21359 Bacteri
20	36	81.8	663	8	AD543005 Bacteri
21	36	81.8	698	7	ADN25464 Hyperther
22	35	79.5	576	6	AD535684 Actinoda
23	35	79.5	739	6	ABU25261 Protein e
24	35	79.5	776	7	ABO68824 Pseudomon
25	34	77.3	10	4	AA964400 Human com

26	34	77.3	472	5	ABP66302	ADP66302 Bifidobac
27	34	77.3	526	8	ABM84012	ADm84012 Human dia
28	34	77.3	567	8	ADH35316	ADh35316 ENZM prot
29	34	77.3	578	8	ABM84011	ADm84011 Human dia
30	34	77.3	608	6	ADA05834	ADA05834 Human NOV
31	34	77.3	608	6	ABU37938	ABU37938 NOXA prot
32	34	77.3	608	7	ADCI4219	ADCI4219 Human enz
33	34	77.3	608	8	ADN63246	ADn63246 Human NOV
34	34	77.3	608	8	ADN62997	ADn62997 Human NOV
35	34	77.3	635	8	ABM84010	ABm84010 Human dia
36	34	77.3	640	3	AA780296	AA780296 Human mt
37	34	77.3	640	6	ADA05832	ADA05832 Human NOV
38	34	77.3	640	8	ADN62995	ADn62995 Human NOV
39	34	77.3	640	8	ADQ89358	ADq89358 Antagonis
40	34	77.3	640	8	ABM81537	ABm81537 Tumour-as
41	34	77.3	640	8	ADP23430	ADp23430 PRO polyP
42	34	77.3	663	4	ABBS9746	ABbs9746 Drosophi
43	34	77.3	693	8	ADN22619	ADn22619 Bacteri
44	34	77.3	862	6	ABU29553	ABu29553 Protein e
45	34	77.3	879	7	ADC95052	ADc95052 E. faeciu

## ALIGNMENTS

## RESULT 1

ABW01180 standard; peptide: 12 AA.

ABW01180;

15-JAN-2004 (first entry)

Saccharomyces cerevisiae motif IV peptide.

Modulator of translation termination; MTT1; helicase B; antiviral;  
therapy; HCSB; nonsense mutation; yeast.

Saccharomyces cerevisiae.

Location/Qualifiers

Misc-difference 1 /label= Unknown /note= "Xaa may be any amino acid"

Misc-difference 2 /label= Unknown /note= "Xaa may be any amino acid"

Misc-difference 8 /label= Unknown /note= "Xaa may be any amino acid"

Misc-difference 9 /label= Unknown /note= "Xaa may be any amino acid"

Misc-difference 9 /label= Unknown /note= "Xaa may be any amino acid"

US6630294-B1.

07-OCT-2003.

22-JUL-1999; 99US-00359268.

22-JUL-1998; 98US-0093685P.

(UNNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

Pelitz S, Czaplinski K, Dimman JD;

WPI; 2003-810549/76.

Identifying an agent that increases nonsense suppression, for antiviral therapy, by contacting modulator of translation termination (MTT1) in Saccharomyces cerevisiae with a test agent, and detecting specific binding to MTT1.

PS Disclosure; Col 43-44; 0pp; English.

XX The invention relates to a method of identifying an agent that increases  
 CC nonsense suppression, by contacting modulator of translation termination  
 CC (MTTI) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.  
 CC The method is useful for identifying compositions or agents which  
 CC increase nonsense suppression. The invention may also be used for  
 CC antiviral therapy and for suppression of pathological nonsense mutations.  
 CC The present sequence is Saccharomyces cerevisiae motif peptide

XX Sequence 12 AA;  
 SQ

Query Match 86.4%; Score 38; DB 7; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.13;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
 Db 3 ILAGDXXQLP 12

RESULT 2  
 ADP4106  
 ID ADP4106 standard; peptide; 12 AA.  
 AC ADP4106;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Yeast translation termination modulation protein motif IV.  
 XX  
 KW gene therapy; translation termination; RNA helicase; MTTI;  
 KW frameshift frequency; aberrant transcript degradation;  
 KW peptidyl transferase modulation; beta-thalassemia; beta-globin;  
 KW Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
 KW Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;  
 KW Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic fibrosis;  
 KW Kidney Stone; Familial hypercholesterolaemia; Retinitis Pigmentosa;  
 KW Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; yeast.  
 KW  
 XX Saccharomyces cerevisiae.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1..2  
 FT Misc-difference 8..9  
 FT Misc-difference /note= "Any amino acid"  
 FT Misc-difference /note= "Any amino acid"  
 XX  
 XX US2004115787-A1.  
 XX  
 XX 17-JUN-2004.  
 XX  
 XX 28-AUG-2003; 2003US-00652334.  
 XX  
 XX 22-JUL-1998; 98US-0093685P.  
 XX 22-JUL-1999; 99US-00359258.  
 XX  
 XX (PELTZ S.  
 XX (CZAP/) CZAPLINSKI K.  
 XX (DINM/) DINMAN J D.  
 XX  
 XX PELTZ S, Czaplinski K, Dinman JD;  
 XX  
 XX WPI; 2004-449400/42.  
 XX  
 XX Identifying a test composition or agent that modulates the efficiency of  
 XX translation termination comprises contacting the MTTI with the test  
 XX composition or agent, and determining if the test composition or agent  
 XX inhibits the MTTI.  
 XX  
 XX Claim 36; SEQ ID NO 4; 41pp; English.

CC The invention relates to a method of identifying a test composition that  
 CC modulates the efficiency of translation termination comprising contacting  
 CC the RNA helicase MTTI with a composition or agent under conditions  
 CC permitting binding between the MTTI and the composition, detecting  
 CC specific binding of the test composition or agent to the MTTI, and  
 CC determining if the test composition or agent inhibits the MTTI. The  
 CC composition and method are useful for modulating the fidelity of  
 CC translation termination or for identifying agents that affect the  
 CC functional activity of mRNAs by altering frameshift frequency, permit  
 CC monitoring of a termination event, promote degradation of aberrant  
 CC transcripts, and provide modulators (inhibitors/activators) of peptidyl  
 CC transferase activity during initiation, elongation, termination and mRNA  
 CC degradation of translation. The agents, which may be antagonists or  
 CC agonists, are useful in screening, diagnostic and therapeutic purposes,  
 CC for diseases or conditions resulting from or cause premature translation,  
 CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular  
 CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,  
 CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,  
 CC Hirschsprung disease, Cystic fibrosis, Kidney Stones, Familial  
 CC hypercholesterolaemia, Retinitis Pigmentosa, or Neurofibromatosis,  
 CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents  
 CC the amino acid sequence of the yeast translation termination modulation  
 CC protein motif IV.

XX Sequence 12 AA;  
 SQ

Query Match 86.4%; Score 38; DB 8; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.13;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
 Db 3 ILAGDXXQLP 12

RESULT 3  
 AAY77807  
 ID AAY77807 standard; peptide; 13 AA.  
 AC AAY77807;  
 XX  
 XX 31-MAY-2000 (first entry)  
 XX  
 DE Motif IV comprised in a gene modulating translation termination.  
 XX  
 KW Helicase B; HCSB; MTTI; modulator of translation termination; eRF1; eRF3;  
 KW eukaryotic release factor; peptidyl transferase; beta-thalassemia;  
 KW beta-globin; Duchenne/Becker Muscular Dystrophy; antianemic.  
 KW  
 XX Unidentified.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1..13  
 FT Misc-difference /note= "residues indicated Xaa are unspecified"  
 XX  
 XX WO200005586-A2.  
 XX  
 XX 03-FEB-2000.  
 XX  
 XX 22-JUL-1999; 99WO-US016802.  
 XX 22-JUL-1998; 98US-00120435.  
 XX  
 XX (UYNE-) UNIV NEW JERSEY.  
 XX  
 XX PELTZ S, Czaplinski K, Dinman JD;  
 XX  
 XX WPI; 2000-171458/15.  
 XX  
 XX New multiprotein complex which can modulate peptidyl transferase activity  
 XX during translation, useful to treat diseases associated with peptidyl  
 XX transferase activity e.g. Duchenne/Becker Muscular Dystrophy.

PS Claim 36; Page 79; 89pp; English.

XX The invention provides a new multiprotein complex which can modulate  
 CC peptidyl transferase activity during translation. The complex comprises  
 CC the gene encoding Helicase B (HCSB; renamed MTTL, for Modulator of  
 CC translation Termination) and the conserved proteins known to interact and  
 CC carry out translation termination in eukaryotic cells, peptidyl  
 CC eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to  
 CC modulate peptidyl transferase activity during translation in a cell. It  
 CC can be administered therapeutically combined with a carrier in  
 CC pharmaceutical compositions to treat diseases associated with peptidyl  
 CC transferase activity, especially diseases resulting from a nonsense or  
 CC frameshift mutation e.g. beta-thalassemia, beta-globin, Duchenne/Becker  
 CC Muscular Dystrophy etc. It can be used to identify disease conditions  
 CC involving a defect in the complex, by transfecting cells with encoding  
 CC nucleic acid and determining the proportion of defective complex before  
 CC and after transfection. It is also useful to screen for drugs involved in  
 CC peptidyl transferase activity during translation, inhibiting the  
 CC interaction between MTTL and eRF3 or involved in enhancing translation  
 CC termination. Vectors comprising polynucleotides encoding the complex (or  
 CC antisense sequences) can be constructed and introduced into cells to  
 CC interfere with complex expression and so modulate the efficiency of  
 CC translation termination of mRNA and/or degradation of aberrant  
 CC transcripts in a cell. Agents binding to the complex can be identified  
 CC and included in therapeutic compositions useful as above, and/or used to  
 CC modulate peptidyl transferase activity during translation in cells. They  
 CC are also useful to modulate the efficiency of translation termination of  
 CC mRNA at a nonsense codon and/or promote degradation of aberrant  
 CC transcripts in cells. The method can be used to identify agents/  
 CC compositions modulating binding to MTTL, useful to identify genes.  
 CC Sequences AA577804-812 represent motifs I-IX comprised in the genes of  
 CC interest, used for modulating translation termination

XX Sequence 13 AA;

Query Match 86.4%; Score 38; DB 3; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.15;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ILADXXQLP 12  
 |||||  
 Db 3 ILADXXQLP 12

RESULT 4  
 ADS21255  
 ID ADS21255 standard; protein; 611 AA.

XX ADS21255;

XX 02-DEC-2004 (first entry)

XX Bacterial polypeptide #10288.

XX Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.  
 PA (SLAT/) SLATER S C.  
 PA (CHEN/) CHEN X.  
 PA (GOLD/) GOLDMAN B S.  
 FI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
 DR WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.

PS Claim 1; SEQ ID NO 10288; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition. Improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX Sequence 611 AA;

Query Match 86.4%; Score 38; DB 8; Length 611;  
 Best Local Similarity 80.0%; Pred. No. 10;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ILADXXQLP 12  
 |||||  
 Db 376 ILADXXQLP 385

RESULT 5  
 ADS43096  
 ID ADS43096 standard; protein; 642 AA.

XX ADS43096;

XX 02-DEC-2004 (first entry)

XX Bacterial polypeptide #21526.

XX Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.

XX Bacteria.

XX US2003233675-A1.

PB		18-DEC-2003.	
XX	PF	20-FEB-2003; 2003US-00369493.	
XX	PR	21-FEB-2002; 2002US-0360039P.	
XX	PA	(CNOV/) CNO Y.	
XX	PA	(HINK/) HINKLE G J.	
XX	PA	(SLAT/) SLATER S C.	
XX	PA	(CHEN/) CHEN X.	
XX	PI	(GOLD/) GOLDMAN B S.	
XX	DR	Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;	
XX	WP1	; 2004-061375/06.	
PT	PT	New recombinant DNA construct comprising a promoter positioned to provide	
PT	PT	for expression of a polynucleotide encoding a polypeptide from a	
PT	PT	microbial source, useful for producing plants with improved properties.	
PS	Claim 1;	SEQ ID NO 21526; 122pp; English.	
XX	CC	The invention relates to a recombinant DNA construct comprising a	
CC	CC	promoter functional in a plant cell, where the promoter is positioned to	
CC	CC	provide for expression of a polynucleotide encoding a polypeptide from a	
CC	CC	microbial source. The invention also relates to a transformed plant	
CC	CC	comprising the recombinant DNA construct and a method of producing a	
CC	CC	transformed plant having an improved property. The plant is a crop plant	
CC	CC	such as maize or soybean. The method of producing a transformed plant	
CC	CC	having an improved property comprises transforming a plant with the	
CC	CC	recombinant DNA construct and growing the transformed plant, where the	
CC	CC	polynucleotide or polypeptide is useful for improving plant properties.	
CC	CC	The recombinant DNA construct is useful for producing plants with	
CC	CC	improved plant properties, e.g. improved cold, heat or drought tolerance,	
CC	CC	tolerance to herbicides, extreme osmotic conditions, pathogens or pests,	
CC	CC	increased resistance to plant disease, better growth rate by modification	
CC	CC	of the cell cycle pathway with plant growth regulators, increased rate of	
CC	CC	homologous recombination, modified seed oil or protein yield and/or	
CC	CC	content, improved yield by modification of carbohydrate, nitrogen or	
CC	CC	phosphorus use and/or uptake, by modification of photosynthesis or by	
CC	CC	providing improved plant growth and development under at least one stress	
CC	CC	condition, improved lignin production or improved galactomannan	
CC	CC	production. This sequence represents a bacterial polypeptide used in the	
CC	CC	scope of the invention. Note: The sequence data for this patent did not	
CC	CC	form part of the printed specification but was obtained in electronic	
CC	CC	format from USPTO at seqdata.uspto.gov/sequence.html.	
XX	SO	Sequence 642 AA;	
	Query Match	86.4%; Score 36; DB 8; Length 642;	
	Best Local Similarity	80.0%; Pred. No. 11;	
	Matches	8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
DY		3 ILAGDXXQLP 12	
DB		390 ILAGDHRLP 399	
RESULT 6			
ID	ADSA42927	standard; protein; 648 AA.	
AC	ADS42927;		
XX	DT	02-DEC-2004 (first entry)	
XX	DE	Bacterial polypeptide #21357.	
XX	REcombinant DNA construct; transformed plant; improved plant property;		
KW	cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;		
KW	pathogen tolerance; pest tolerance; plant disease resistance;		
KW	cell cycle pathway modification; plant growth regulator;		
KW	homologous recombination; seed oil yield; protein yield; carbohydrate;		

XX	nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX	bacterial polypeptide.
XX	Bacteria.
XX	US200323675-A1.
XX	18-DEC-2003.
XX	20-FEB-2003; 2003US-00369493.
XX	21-FEB-2002; 2002US-0360039P.
XX	(CAOY/) CAO Y.
XX	(HINK/) HINKLE G J.
XX	(SLAT/) SLATER S C.
XX	(CHEN/) CHEN X.
XX	(GOLD/) GOLDMAN B S.
XX	Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX	WPI; 2004-061375/06.
XX	New recombinant DNA construct comprising a promoter positioned to provide
XX	for expression of a polynucleotide encoding a polypeptide from a
XX	microbial source, useful for producing plants with improved properties.
XX	Claim 1; SEQ ID NO 21357; 122pp; English.
XX	The invention relates to a recombinant DNA construct comprising a
XX	promoter functional in a plant cell, where the promoter is positioned to
XX	provide for expression of a polynucleotide encoding a polypeptide from a
XX	microbial source. The invention also relates to a transformed plant
XX	comprising the recombinant DNA construct and a method of producing a
XX	transformed plant having an improved property. The plant is a crop plant
XX	such as maize or soybean. The method of producing a transformed plant
XX	having an improved property comprises transforming a plant with the
XX	recombinant DNA construct and growing the transformed plant, where the
XX	polynucleotide or polypeptide is useful for improving plant properties.
XX	The recombinant DNA construct is useful for producing plants with
XX	improved plant properties, e.g. improved cold, heat or drought tolerance,
XX	tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX	increased resistance to plant disease, better growth rate by modification
XX	of the cell cycle pathway with plant growth regulators, increased rate of
XX	homologous recombination, modified seed oil or protein yield and/or
XX	content, improved yield by modification of carbohydrate, nitrogen or
XX	phosphorus use and/or uptake, by modification of photosynthesis or by
XX	providing improved plant growth and development under at least one stress
XX	condition, improved lignin production or improved galactomannan
XX	production. This sequence represents a bacterial polypeptide used in the
XX	scope of the invention. Note: The sequence data for this patent did not
XX	form part of the printed specification but was obtained in electronic
XX	format from USPTO at seqdata.uspto.gov/sequence.html.
XX	Sequence 648 AA;
XX	Query Match 86.4%; Score 38; DB 8; Length 648;
XX	Best Local Similarity 80.0%; Pred. No. 11;
XX	Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX	3 ILAGDXXQLP 12
XX	
XX	398 ILAGDHRQLP 407
XX	
XX	RESULT 7
XX	ADS41904
XX	ADS41904 standard; protein; 648 AA.
XX	ADS41904;
XX	02-DEC-2004 (first entry)
XX	DT



DE Bacterial polypeptide #20334.

XX Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.

OS Bacteria.

FN US2003233675-A1.

PD 18-DEC-2003.

PP 20-FEB-2003; 2003US-00369493.

PR 21-FEB-2002; 2002US-0360039P.

PA (CAOY/) CAO Y.

PA (HINKLE/) HINKLE G. J.

PA (SLATER/) SLATER S. C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B. S.

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

DR WPI; 2004-061375/06.

PT New recombinant DNA construct comprising a promoter positioned to provide

PT for expression of a polynucleotide encoding a polypeptide from a

PT microbial source, useful for producing plants with improved properties.

PS Claim 1; SEQ ID NO 20334; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition. Improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 648 AA;

Query Match 86.4%; Score 38; DB 8; Length 648;

Best Local Similarity 80.0%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

3 ILAGDXXQLP 12

405 ILAGDXXQLP 414

RESULT 8

AAB62031 ID AAB62031 standard; protein; 655 AA.

XX AAB62031;

AC 14-MAY-2001 (first entry)

DT Recombinant P. furiosus helicase dna2.

DE MCM, minichromosome maintenance protein; archaeal polypeptide; PCNA;

KW RFC-P38; RFC-P55; RPA; CDC6; PEN-1; dUTPase; ligase; helicase dna2; PCR;

KW nucleic acid amplification; polymerase chain reaction.

OS Pyrococcus furiosus.

FN WO200109347-A2.

PD 08-FEB-2001.

PP 28-JUL-2000; 2000WO-US020532.

PR 30-JUL-1999; 99US-0146580P.

PA (STRA-) STRATAGENE.

PI Hogrefe HH, Cline JM, Hansen CJ, Borne MC;

DR WPI; 2001-182959/18.

DR N-PSDB; AAF57035.

PT Composition for improving nucleic acid polymerase reactions, useful e.g.

PT in synthesis or amplification, contains at least one archaeal accessory

PT protein.

PS Claim 182; Fig 35; 147pp; English.

XX The invention provides a composition (A) for enhancing nucleic acid  
 CC polymerase reactions that comprises an archaeal MCM (minichromosome  
 CC maintenance protein) and at least one of the archaeal polypeptides (PCNA,  
 CC RFC-P38 or -P55, RPA, CDC6, PEN-1, dUTPase, ligase, helicase dna2, or  
 CC helicases 2-8). (A) And similar compositions containing different  
 CC combinations of accessory proteins, are used to improve performance of  
 CC synthesis, amplification, mutagenizing, labeling and detecting reactions,  
 CC e.g. for gene characterization, cloning, detection of allelic variants,  
 CC diagnosis and screening for disease, particularly where done by  
 CC polymerase chain reaction (PCR). Some of the proteins also stabilize  
 CC duplexes during polymerase reactions or improve exonuclease reactions,  
 CC for example RPA also improves specificity of nucleic acid/protein  
 CC interaction and PCNA improves polymerase-mediated repair processes and  
 CC hybridization reactions. Nucleic acids encoding the archaeal polypeptides  
 CC are used for recombinant production of proteins, and fragments of the  
 CC nucleic acid as probes and primers for screening related sequences. The  
 CC accessory proteins increase accuracy and efficiency of polymerase  
 CC reactions, allow use of lower denaturation and extension temperatures  
 CC (possibly isothermal processing), and improve synthesis of long targets.  
 CC The present sequence represents a P. furiosus recombinant helicase dna2

XX Sequence 655 AA;

Query Match 86.4%; Score 38; DB 4; Length 655;

Best Local Similarity 80.0%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

3 ILAGDXXQLP 12

405 ILAGDXXQLP 414

RESULT 9

ADN18615 ID ADN18615 standard; protein; 656 AA.

AC ADN18615;

XX 02-DEC-2004 (first entry)  
 DT Bacterial polypeptide #1268.  
 DE  
 XX Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.  
 KM  
 XX Bacteria.  
 OS  
 XX US2003233675-A1.  
 PN  
 XX 18-DEC-2003.  
 PD  
 XX 20-FEB-2003; 2003US-00369493.  
 PF  
 XX 21-FEB-2002; 2002US-0360039P.  
 PR  
 XX (CAOY/) CAO Y.  
 PA (HINK/) HINKLE G J.  
 PA (SLAT/) SLATER S C.  
 PA (CHEN/) CHEN X.  
 PA (GOLD/) GOLDMAN B S.  
 XX  
 P1 Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
 XX  
 DR WPI; 2004-061375/06.  
 DR  
 XX  
 PT New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.  
 PT  
 XX  
 PS Claim 1; SEQ ID NO 1268; 122pp; English.  
 PS  
 XX The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition, improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at seqdata.uspto.gov/sequence.html.  
 CC  
 XX  
 SO Sequence 656 AA;  
 SO  
 Query Match 86.4%; Score 38; DB 8; Length 656;  
 Best Local Similarity 80.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 ILAGDXXQLP 12  
 DB 405 ILAGDHRQLP 414

RESULT 10  
 AA026745  
 ID AA026745 standard; protein; 988 AA.  
 XX  
 AC AA026745;  
 XX  
 DT 27-MAR-2003 (first entry)  
 XX  
 DE 988-mer rat protein sequence, SEQ ID NO 2.  
 XX  
 KW Antiviral; gene therapy; transcription controller; antiretrovirus agent;  
 KW tannic acid responsive element-coupling protein; nerve growth;  
 KW maintenance controlling agent; retrovirus; rat.  
 KM  
 XX Rattus sp.  
 OS  
 XX JP2002262878-A.  
 PN  
 XX 17-SEP-2002.  
 PD  
 XX 08-MAR-2001; 2001JP-00065803.  
 PF  
 XX 08-MAR-2001; 2001JP-00065803.  
 PR  
 XX (TANU/) TANUMA Y.  
 PA (SUMU) SUMITOMO SEIYAKU KK.  
 PA (HUMA-) ZH HUMAN SCI SHINKO ZAIDAN.  
 XX  
 XX WPI; 2003-132123/13.  
 DR N-PSDB; AAL54290.  
 DR  
 XX  
 PT A transcription controller, useful for inhibiting the growth of a  
 PT retrovirus, comprises a sequence encoding a tannic acid responsive  
 PT element-coupling protein.  
 PT  
 XX  
 PS Claim 2; Fig 2; 28pp; Japanese.  
 PS  
 XX The invention relates to a transcription controller comprising a nucleic  
 CC acid encoding a tannic acid responsive element-coupling protein. The  
 CC transcription controller is useful in an antiretrovirus agent and nerve  
 CC growth/maintenance controlling agent for growth inhibition of a  
 CC retrovirus. The protein can be used for developing a new low molecular  
 CC compound controlling transcription control. This sequence represents a  
 CC rat protein sequence relating to the transcription controller of the  
 CC invention.  
 CC  
 XX  
 SO Sequence 988 AA;  
 SO  
 Query Match 86.4%; Score 38; DB 6; Length 988;  
 Best Local Similarity 80.0%; Pred. No. 18;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 ILAGDXXQLP 12  
 DB 395 ILAGDHRQLP 404  
 RESULT 11  
 AA077816  
 ID AA077816 standard; peptide; 414 AA.  
 XX  
 AC AA077816;  
 XX  
 DT 31-MAY-2000 (first entry)  
 DT  
 XX Yeast Dlp1 protein fragment.  
 DE  
 XX Helicase B; HCSB; WTT1; modulator of translation termination; eRF1; eRF3;  
 KW eukaryotic release factor; peptidyl transferase; beta-thalassemia;  
 KW beta-globin; Duchenne/Becker Muscular Dystrophy; antianemic; yeast;  
 KW helicase; Dlp1.

XX	08	Saccharomyces cerevisiae.
XX	PN	WO200005586-A2.
XX	PD	03-FEB-2000.
XX	PP	22-JUL-1999; 99WO-US016802.
XX	PR	22-JUL-1999; 98US-00120435.
XX	PA	(UNRE-) UNIV NEW JERSEY.
XX	PI	Peltz S, Czaplinski K, Dinman JD;
XX	DR	WPI; 2000-171458/15.
PT	PT	New multiprotein complex which can modulate peptidyl transferase activity during translation, useful to treat diseases associated with peptidyl transferase activity e.g. Duchenne/Becker Muscular Dystrophy.
PS	PS	Example 1; Fig 1; 89pp; English.
XX	XX	The invention provides a new multiprotein complex which can modulate peptidyl transferase activity during translation. The complex comprises the gene encoding Helicase B (HCSB; renamed MTT1, for Modulator of Translation Termination) and the conserved proteins known to interact and carry out translation termination in eukaryotic cells, peptidyl eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to modulate peptidyl transferase activity during translation in a cell. It can be administered therapeutically combined with a carrier in pharmaceutical compositions to treat diseases associated with peptidyl transferase activity, especially diseases resulting from a nonsense or frameshift mutation e.g. beta-thalassemia, beta-globin, Duchenne/Becker Muscular Dystrophy etc. It can be used to identify disease conditions involving a defect in the complex, by transfecting cells with encoding nucleic acid and determining the proportion of defective complex before and after transfection. It is also useful to screen for drugs involved in peptidyl transferase activity during translation, inhibiting the interaction between MTT1 and eRF3 or involved in enhancing translation termination. Vectors comprising polynucleotides encoding the complex (or antisense sequences) can be constructed and introduced into cells to interfere with complex expression and so modulate the efficiency of translation termination of mRNA and/or degradation of aberrant transcripts in a cell. Agents binding to the complex can be identified and included in therapeutic compositions useful as above, and/or used to modulate peptidyl transferase activity during translation in cells. They are also useful to modulate the efficiency of translation termination of mRNA at a nonsense codon and/or promote degradation of aberrant transcripts in cells. The method can be used to identify agents/compositions modulating binding to MTT1, useful to identify genes. Sequences AAYV7813-817 represent protein fragments from yeast superfamily group I helicases
XX	XX	Sequence 414 AA;
XX	XX	Query Match 84.1%; Score 37; DB 3; Length 414;
XX	XX	Best local similarity 70.0%; Pred. No. 11;
XX	XX	Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0
XX	XX	3 ILAGDXXKQLP 12
XX	XX	:
XX	XX	193 VLAGDNKQLP 202
XX	XX	RESULT 12
XX	XX	ABW01204
XX	XX	ABW01204 standard; protein; 414 AA.
XX	XX	ABW01204;
XX	XX	15-JAN-2004 (first entry)
XX	XX	

DE	Saccharomyces cerevisiae	Dipl protein.
XX		
KM	Modulator of translation termination; MTT1; helicase B; antiviral;	
KW	therapy; HCSB; nonsense mutation; yeast.	
XX		
OS	Saccharomyces cerevisiae.	
XX		
XX	US6630294-B1.	
PN		
PD	07-OCT-2003.	
XX		
XX	22-JUL-1999;	99US-00359268.
PP		
XX	22-JUL-1998;	98US-0093685P.
PR		
PA	(UTNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.	
XX		
PI	Peltz S, Czaplinski K, Dلمان JD;	
XX		
DR	WPI; 2003-810549/76.	
XX		
PT	Identifying an agent that increases nonsense suppression, for antiviral	
PT	therapy, by contacting modulator of translation termination (Mtl1) in	
PT	Saccharomyces cerevisiae with a test agent, and detecting specific	
PT	binding to Mtl1.	
PS	Disclosure; Col 55-58; Opp; English.	
XX		
CC	The invention relates to a method of identifying an agent that increases	
CC	nonsense suppression, by contacting modulator of translation termination	
CC	(MTT1) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.	
CC	The method is useful for identifying compositions or agents which	
CC	increase nonsense suppression. The invention may also be used for	
CC	antiviral therapy and for suppression of pathological nonsense mutations.	
CC	The present sequence is Saccharomyces cerevisiae Dipl protein	
XX		
SQ	Sequence 414 AA;	
	Query Match	84.1%; Score 37; DB 7; Length 414;
	Best Local Similarity	70.0%; Pred. No. 11;
	Matches	7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY	3 ILADXXQLP 12	
	:	
	193 VLADNRQLP 202	
DB		
	RESULT 13	
ADP44130		
ID	ADP44130 standard; protein; 414 AA.	
XX		
AC	ADP44130;	
XX		
DT	18-NOV-2004 (first entry)	
XX		
DS	Yeast helicase Dipl.	
XX		
KM	gene therapy; translation termination; RNA helicase; MTT1;	
KW	frameshift frequency; aberrant transcript degradation;	
KW	peptidyl transferase modulation; beta-thalassemia; beta-globin;	
KW	Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;	
KW	Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;	
KW	Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic fibrosis;	
KW	Kidney Stone; Familial hypercholesterolaemia; Retinitis Pigmentosa;	
XX	Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; yeast; enzyme.	
OS	Saccharomyces cerevisiae.	
XX		
PN	US2004115787-A1.	
XX		
PD	17-JUN-2004.	
XX		
PF	28-AUG-2003; 2003US-00652334.	



KM cell cycle pathway modification; plant growth regulator;  
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KM bacterial polypeptide.

XX Bacteria.

PN US2003233675-A1.

PD 18-DEC-2003.

PF 20-FEB-2003; 2003US-00369493.

PR 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.  
 PA (HINK/) HINKLE G J.  
 PA (SLAT/) SLATER S C.  
 PA (CHEN/) CHEN X.  
 PA (GOLD/) GOLDMAN B S.

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

DR WPI; 2004-061375/06.

PT New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.

PS Claim 1; SEQ ID NO 21645; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition, improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 653 AA;

Query Match 84.1%; Score 37; DB 8; Length 653;  
 Best Local Similarity 70.0%; Pred. No. 19;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 ILAGDXXQLP 12  
 :|||||  
 Db 405 VIAGDHRQLP 414

Search completed: April 18, 2005, 08:03:46  
 Job time : 61.8483 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 07:37:21 ; Search time 14.9781 Seconds  
(without alignments)  
59.807 Million cell updates/sec

Title: US-10-652-334-4

Sequence: 1 XXIIAGDXQLP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCCTS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	86.4	12	US-09-359-268A-4	Sequence 4, Appl1
2	38	86.4	993	US-09-538-092-1100	Sequence 1100, Ap
3	37	84.1	414	US-09-359-268A-28	Sequence 28, Appl
4	37	84.1	683	US-09-538-092-483	Sequence 483, App
5	35	79.5	426	US-09-248-796A-15170	Sequence 15170, A
6	35	79.5	576	US-09-328-352-6971	Sequence 6971, Ap
7	35	79.5	776	US-09-252-991A-17570	Sequence 17570, A
8	34	77.3	513	US-09-902-540-13564	Sequence 13564, A
9	34	77.3	879	US-09-107-532A-4679	Sequence 4679, Ap
10	33	75.0	102	US-09-270-767-58044	Sequence 58044, A
11	33	75.0	152	US-09-270-767-42726	Sequence 42726, A
12	33	75.0	723	US-09-425-335-5	Sequence 5, Appl1
13	33	75.0	805	US-09-425-335-6	Sequence 6, Appl1
14	32	72.7	282	US-09-248-796A-50596	Sequence 20596, A
15	32	72.7	287	US-09-328-352-6505	Sequence 6205, Ap
16	32	72.7	295	US-09-198-452A-801	Sequence 801, App
17	32	72.7	366	US-09-359-268A-27	Sequence 27, Appl
18	32	72.7	366	US-09-438-185A-754	Sequence 754, App
19	32	72.7	677	US-09-425-335-3	Sequence 3, Appl1
20	32	72.7	689	US-09-425-335-2	Sequence 2, Appl1
21	32	72.7	751	US-09-036-987A-24	Sequence 24, Appl
22	32	72.7	751	US-09-370-700-24	Sequence 24, Appl
23	32	72.7	751	US-09-603-207-24	Sequence 24, Appl
24	31	70.5	108	US-08-560-398-6	Sequence 6, Appl1
25	31	70.5	239	US-09-902-540-12278	Sequence 12278, A
26	31	70.5	428	US-09-489-039A-12939	Sequence 12939, A
27	31	70.5	453	US-09-540-236-3464	Sequence 3464, Ap

28	31	70.5	472	US-09-359-268A-26	Sequence 26, Appl
29	31	70.5	508	US-09-252-991A-32269	Sequence 32269, A
30	31	70.5	788	US-09-583-110-3939	Sequence 3939, Ap
31	31	70.5	856	US-09-107-433-3534	Sequence 3534, Ap
32	30	68.2	202	US-09-252-991A-22316	Sequence 22316, A
33	30	68.2	407	US-09-634-238-240	Sequence 240, App
34	30	68.2	422	US-09-248-796A-16512	Sequence 16512, A
35	30	68.2	514	US-09-489-039A-10028	Sequence 10028, A
36	30	68.2	539	US-09-134-000C-4363	Sequence 4363, Ap
37	30	68.2	556	US-09-425-335-7	Sequence 7, Appl1
38	30	68.2	648	US-09-902-540-13835	Sequence 13835, A
39	30	68.2	770	US-09-489-039A-7872	Sequence 7872, Ap
40	30	68.2	857	US-09-425-335-4	Sequence 4, Appl1
41	29	65.9	118	US-09-949-016-7744	Sequence 7744, Ap
42	29	65.9	154	US-09-270-767-59910	Sequence 59910, A
43	29	65.9	165	US-09-270-767-37774	Sequence 37774, A
44	29	65.9	165	US-09-270-767-52991	Sequence 52991, A
45	29	65.9	218	US-10-026-045-2	Sequence 2, Appl1

#### ALIGNMENTS

```

RESULT 1
US-09-359-268A-4
; Sequence 4, Application US/09359268A
; Patent No. 6630294
;
GENERAL INFORMATION:
; APPLICANT: Peltz, Stuart
; APPLICANT: Cepialinski, Kevin
; APPLICANT: Dimman, Jonathan D.
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
; FILE REFERENCE: 601-1-85N
; CURRENT APPLICATION NUMBER: US/09/359,268A
; CURRENT FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,685
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURES:
; OTHER INFORMATION: Xaa = any amino acid
US-09-359-268A-4.

Query Match      86.4% Score 38; DB 4; Length 12;
Query Local Similarity 100.0%; Pred. No. 0.031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 ILAGDXQLP 12
        |||||
        3 ILAGDXQLP 12

DB
US-09-538-092-1100
; Sequence 1100, Application US/09538092
; Patent No. 675314
;
GENERAL INFORMATION:
; APPLICANT: Glac, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01

```

NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: CurataseqFormatter Version 0.9  
SEQ ID NO 1100  
LENGTH: 993  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (0) ... (0)  
OTHER INFORMATION: Polypeptide Accession Number P38935  
US-09-538-092-1100

Query Match 86.4%; Score 38; DB 4; Length 993;  
Best Local Similarity 80.0%; Pred. No. 4.2;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ILAGDXXQLP 12  
:|||||  
Db 396 ILAGDHRQLP 405

RESULT 3  
US-09-359-268A-28  
Sequence 28, Application US/09359268A  
Patent No. 6630294  
GENERAL INFORMATION:  
APPLICANT: Peltz, Stuart  
APPLICANT: Czaplinski, Kevin  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/09/359,268A  
CURRENT FILING DATE: 1998-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
PRIOR FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 28  
LENGTH: 414  
TYPE: PRT  
ORGANISM: saccharomyces cerevisiae  
US-09-359-268A-28

Query Match 84.1%; Score 37; DB 4; Length 414;  
Best Local Similarity 70.0%; Pred. No. 2.7;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 ILAGDXXQLP 12  
:|||||  
Db 193 VLAGDNKQLP 202

RESULT 4  
US-09-538-092-483  
Sequence 483, Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:  
APPLICANT: Glot, Loic  
APPLICANT: Mansfield, Traci A.  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: CurataseqFormatter Version 0.9  
SEQ ID NO 483  
LENGTH: 683

TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (0) ... (0)  
OTHER INFORMATION: Polypeptide Accession Number YKL017C  
US-09-538-092-483

Query Match 84.1%; Score 37; DB 4; Length 683;  
Best Local Similarity 70.0%; Pred. No. 4.6;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 ILAGDXXQLP 12  
:|||||  
Db 420 VLAGDNKQLP 429

RESULT 5  
US-09-248-796A-15170  
Sequence 15170, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 15170  
LENGTH: 426  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-15170

Query Match 79.5%; Score 35; DB 4; Length 426;  
Best Local Similarity 60.0%; Pred. No. 7.8;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 ILAGDXXQLP 12  
:|||||  
Db 142 VLAGDNKQLP 151

RESULT 6  
US-09-328-352-6971  
Sequence 6971, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 6971  
LENGTH: 576  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-6971

Query Match 79.5%; Score 35; DB 4; Length 576;  
Best Local Similarity 70.0%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 ILAGDXXQLP 12  
:|||||  
Db 140 IVAGDPFQLP 149



RESULT 7  
US-09-252-991A-17570  
Sequence 17570, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 17570  
LENGTH: 776  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17570

Query Match 79.5%; Score 35; DB 4; Length 776;  
Best Local Similarity 60.0%; Pred. No. 15;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 ILAGDXQLP 12  
:|||||:  
Db 399 ILAGDATELP 408

RESULT 8  
US-09-902-540-13564  
Sequence 13564, Application US/09902540  
Patent No. 6833447  
GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Wiegand, Roger C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(15849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 13564  
LENGTH: 513  
TYPE: PRT  
ORGANISM: Myxococcus xanthus  
US-09-902-540-13564

Query Match 77.3%; Score 34; DB 4; Length 513;  
Best Local Similarity 60.0%; Pred. No. 15;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 ILAGDXQLP 12  
:|||||:  
Db 77 VIAGDPKQLP 86

RESULT 9  
US-09-107-532A-4679  
Sequence 4679, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02154

COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:  
NAME: Ariandello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 4679:

SEQUENCE CHARACTERISTICS:  
LENGTH: 879 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULAR TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium

FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...879  
SEQUENCE DESCRIPTION: SEQ ID NO: 4679

US-09-107-532A-4679

Query Match 77.3%; Score 34; DB 4; Length 879;  
Best Local Similarity 70.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 ILAGDXQLP 12  
:|||||:  
Db 487 ILVGDPKQLP 496

RESULT 10  
US-09-270-767-58044  
Sequence 58044, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 58044  
LENGTH: 102  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-270-767-58044

Query Match 75.0%; Score 33; DB 4; Length 102;  
Best Local Similarity 60.0%; Pred. No. 4.5;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
: ||| |||  
Db 36 VLVGDMQQLP 45

## RESULT 11

US-09-270-767-42726  
; Sequence 42726, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 42726  
; LENGTH: 152  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-42726

Query Match 75.0%; Score 33; DB 4; Length 152;  
Best Local Similarity 60.0%; Pred. No. 7;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
: ||| |||  
Db 86 VLVGDMQQLP 95

## RESULT 12

US-09-425-335-5  
; Sequence 5, Application US/09425335  
; Patent No. 6518052  
; GENERAL INFORMATION:  
; APPLICANT: WEIMANN, ROBERTO  
; TITLE OF INVENTION: HUMAN HOMOLOGUE OF YEAST HELICASE AND USES THEREOF  
; FILE REFERENCE: db7 sequence  
; CURRENT APPLICATION NUMBER: US/09/425,335  
; CURRENT FILING DATE: 1999-10-22  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 723  
; TYPE: PRT  
; ORGANISM: Yeast homolog  
US-09-425-335-5

Query Match 75.0%; Score 33; DB 4; Length 723;  
Best Local Similarity 60.0%; Pred. No. 40;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
: ||| |||  
Db 370 VLVGDMQQLP 379

## RESULT 13

US-09-425-335-6  
; Sequence 6, Application US/09425335  
; Patent No. 6518052  
; GENERAL INFORMATION:  
; APPLICANT: WEIMANN, ROBERTO  
; TITLE OF INVENTION: HUMAN HOMOLOGUE OF YEAST HELICASE AND USES THEREOF  
; FILE REFERENCE: db7 sequence  
; CURRENT APPLICATION NUMBER: US/09/425,335  
; CURRENT FILING DATE: 1999-10-22  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 805

; TYPE: PRT  
; ORGANISM: Schizosaccharomyces pombe  
US-09-425-335-6

Query Match 75.0%; Score 33; DB 4; Length 805;  
Best Local Similarity 60.0%; Pred. No. 45;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
: ||| |||  
Db 450 VLVGDMQQLP 459

## RESULT 14

US-09-248-796A-20596  
; Sequence 20596, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Kelch Weinstein et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 20596  
; LENGTH: 282  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-20596

Query Match 72.7%; Score 32; DB 4; Length 282;  
Best Local Similarity 50.0%; Pred. No. 24;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
: ||| |||  
Db 146 LLSGDSQVP 155

## RESULT 15

US-09-328-352-6205  
; Sequence 6205, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Brelton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6205  
; LENGTH: 287  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-6205

Query Match 72.7%; Score 32; DB 4; Length 287;  
Best Local Similarity 66.7%; Pred. No. 24;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 LAGDXXQLP 12  
: ||| |||  
Db 220 LSGDALQLP 228

Search completed: April 18, 2005, 08:18:34  
Job time : 15.9781 secs

---

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 18, 2005, 08:06:16 / Search time 42.5693 Seconds  
(without alignments)  
93.693 Million cell updates/sec

Title: US-10-652-334-4  
Perfect score: 44  
Sequence: 1 XXILAGDXXQLP 12

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database : Published Applications\_AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	86.4	12	US-10-652-334-4	Sequence 4, Appl1
2	38	86.4	332	US-10-425-114-38637	Sequence 3837, A
3	38	86.4	404	US-10-425-114-37833	Sequence 37833, A
4	38	86.4	611	US-10-369-493-10288	Sequence 10288, A
5	38	86.4	637	US-10-424-559-233501	Sequence 233501, A
6	38	86.4	639	US-10-425-114-37717	Sequence 37717, A
7	38	86.4	642	US-10-369-493-21526	Sequence 21526, A
8	38	86.4	642	US-10-425-114-37557	Sequence 37557, A
9	38	86.4	648	US-10-369-493-20334	Sequence 20334, A
10	38	86.4	648	US-10-369-493-21357	Sequence 21357, A
11	38	86.4	655	US-10-828-924-80	Sequence 80, Appl1
12	38	86.4	656	US-10-369-493-1268	Sequence 1268, Ap
13	37	84.1	414	US-10-652-334-28	Sequence 28, Appl1

14	37	84.1	626	US-10-437-963-166322	Sequence 166322, A
15	37	84.1	650	US-10-369-493-2884	Sequence 2884, Ap
16	37	84.1	653	US-10-369-493-21645	Sequence 21645, A
17	37	84.1	683	US-10-369-493-22264	Sequence 22264, A
18	37	84.1	830	US-10-369-493-4012	Sequence 4012, Ap
19	36	81.8	663	US-10-369-493-21435	Sequence 21435, A
20	35	79.5	640	US-10-926-543-76	Sequence 76, Appl1
21	35	79.5	739	US-10-282-122A-53185	Sequence 53185, A
22	34	77.3	10	US-09-572-1048-2594	Sequence 2594, A
23	34	77.3	608	US-10-262-511-194	Sequence 194, App
24	34	77.3	640	US-10-262-511-192	Sequence 192, App
25	34	77.3	693	US-10-369-493-5272	Sequence 5272, Ap
26	34	77.3	862	US-10-282-122A-57477	Sequence 57477, A
27	33	75.0	133	US-10-424-599-247410	Sequence 247410, A
28	33	75.0	309	US-10-425-114-39052	Sequence 39052, A
29	33	75.0	312	US-10-425-114-45797	Sequence 45797, A
30	33	75.0	416	US-10-424-599-277834	Sequence 277834, A
31	33	75.0	472	US-10-282-122A-48260	Sequence 48260, A
32	33	75.0	562	US-10-104-047-2003	Sequence 2003, Ap
33	33	75.0	581	US-10-282-122A-65905	Sequence 65905, A
34	33	75.0	592	US-10-282-122A-65055	Sequence 65055, A
35	33	75.0	677	US-10-476-924-9	Sequence 9, Appl1
36	33	75.0	723	US-10-369-493-22031	Sequence 22031, A
37	33	75.0	744	US-10-282-122A-55137	Sequence 55137, A
38	33	75.0	787	US-10-282-122A-72461	Sequence 72461, A
39	33	75.0	798	US-10-282-122A-60935	Sequence 60935, A
40	33	75.0	817	US-10-282-122A-74732	Sequence 74732, A
41	33	75.0	819	US-10-369-493-22550	Sequence 22550, A
42	33	75.0	821	US-10-369-493-18631	Sequence 18631, A
43	33	75.0	828	US-10-144-194A-96	Sequence 96, Appl1
44	33	75.0	1024	US-10-369-493-3447	Sequence 3447, Ap
45	33	75.0	1076	US-09-935-642-12	Sequence 12, Appl1

## ALIGNMENTS

RESULT 1  
US-10-652-334-4  
; Sequence 4, Application US/10652334  
; Publication No. US20040115787A1  
; GENERAL INFORMATION:  
; APPLICANT: Pelcz, Stuart  
; APPLICANT: Czaplinski, Kevin  
; APPLICANT: Dimman, Jonathan D.  
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
; FILE REFERENCE: 601-1-85N  
; CURRENT APPLICATION NUMBER: US/10/652,334  
; PRIOR FILING DATE: 2003-08-28  
; PRIOR APPLICATION NUMBER: US/09/359,268A  
; PRIOR FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 60/093,685  
; PRIOR FILING DATE: 1998-07-22  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
; FEATURES:  
; OTHER INFORMATION: Xaa = any amino acid  
US-10-652-334-4

Query Match 86.4%; Score 38; DB 16; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 3 ILAGDXXQLP 12  
DB 3 ILAGDXXQLP 12

```
RESULT 2
US-10-425-114-38637
; Sequence 38637, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; PRIORITY FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38637
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700241578_FLI.pep
US-10-425-114-38637

Query Match          86.4%; Score 38; DB 15; Length 332;
Best Local Similarity 80.0%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12
Db 82 ILAGDHLQLP 91

RESULT 3
US-10-425-114-37833
; Sequence 37833, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; PRIORITY FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37833
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700618821_FLI.pep
US-10-425-114-37833

Query Match          86.4%; Score 38; DB 15; Length 404;
Best Local Similarity 80.0%; Pred. No. 6.2;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12
Db 154 ILAGDHLQLP 163

RESULT 4
US-10-369-493-10288
; Sequence 10288, Application US/10369493
; Publication No. US20030233675A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIORITY FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10288
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Cytophaga hutchinsonii
US-10-369-493-10288

Query Match          86.4%; Score 38; DB 15; Length 611;
Best Local Similarity 80.0%; Pred. No. 9.6;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12
Db 376 ILAGDHLQLP 385

RESULT 5
US-10-424-599-233501
; Sequence 233501, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; PRIORITY FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 233501
; LENGTH: 637
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(637)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_52878C.1.pep
US-10-424-599-233501

Query Match          86.4%; Score 38; DB 15; Length 637;
Best Local Similarity 80.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12
Db 387 ILAGDHLQLP 396

RESULT 6
US-10-425-114-37717
; Sequence 37717, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
```

APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(5313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 37717  
LENGTH: 639  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: L1B3170-056-H12\_Flt.pep  
US-10-425-114-37717

Query Match 86.4%; Score 38; DB 15; Length 639;  
Best Local Similarity 80.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
|||||  
Db 389 ILAGDHLQLP 398

RESULT 7  
US-10-369-493-21526  
Sequence 21526, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 21526  
LENGTH: 642  
TYPE: PRT  
ORGANISM: Methanobacterium thermoautotrophicum  
US-10-369-493-21526

Query Match 86.4%; Score 38; DB 15; Length 642;  
Best Local Similarity 80.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
|||||  
Db 390 ILAGDHLQLP 399

RESULT 8  
US-10-425-114-37557  
Sequence 37557, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jindong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(5313)B

CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 37557  
LENGTH: 642  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana columbia  
FEATURE:  
OTHER INFORMATION: Clone ID: L1B3175-019-D10\_Flt.pep  
US-10-425-114-37557

Query Match 86.4%; Score 38; DB 15; Length 642;  
Best Local Similarity 80.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
|||||  
Db 393 ILAGDHLQLP 402

RESULT 9  
US-10-369-493-20334  
Sequence 20334, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 20334  
LENGTH: 648  
TYPE: PRT  
ORGANISM: Pyrococcus horikoshii  
US-10-369-493-20334

Query Match 86.4%; Score 38; DB 15; Length 648;  
Best Local Similarity 80.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
|||||  
Db 405 ILAGDHLQLP 414

RESULT 10  
US-10-369-493-21357  
Sequence 21357, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 21357  
LENGTH: 648

TYPE: PRT  
ORGANISM: Archaeoglobus fulgidus  
US-10-369-493-21357

Query Match 86.4%; Score 38; DB 15; Length 648;  
Best Local Similarity 80.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
Db 398 ILAGDHRQLP 407

RESULT 11  
US-10-828-924-80  
Sequence 80, Application US/10828924  
Publication No. US2005003401A1  
GENERAL INFORMATION:  
APPLICANT: STRATAGENE  
TITLE OF INVENTION: PCU REPLICATION ACCESSORY FACTORS AND METHODS OF USE  
FILE REFERENCE: 04121, 0161-00000  
CURRENT APPLICATION NUMBER: US/10/828,924  
PRIOR FILING DATE: 2004-04-20  
PRIOR APPLICATION NUMBER: 60/146,580  
PRIOR FILING DATE: 1999-07-30  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 80  
LENGTH: 655  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Helicase dna2  
US-10-828-924-80

Query Match 86.4%; Score 38; DB 16; Length 655;  
Best Local Similarity 80.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
Db 405 ILAGDHRQLP 414

RESULT 12  
US-10-369-493-1268  
Sequence 1268, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xiaofeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
PRIOR FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 1268  
LENGTH: 656  
TYPE: PRT  
ORGANISM: Pyrococcus horikoshii  
US-10-369-493-1268

Query Match 86.4%; Score 38; DB 15; Length 656;  
Best Local Similarity 80.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12

Db 405 ILAGDHRQLP 414

RESULT 13  
US-10-652-334-28  
Sequence 28, Application US/10652334  
Publication No. US20040115787A1  
GENERAL INFORMATION:  
APPLICANT: Peltz, Stuart  
APPLICANT: Czaplinski, Kevin  
APPLICANT: Dimman, Jonathan D.  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/10/652,334  
CURRENT FILING DATE: 2003-08-28  
PRIOR APPLICATION NUMBER: US/09/359,268A  
PRIOR FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
PRIOR FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 28  
LENGTH: 414  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-10-652-334-28

Query Match 84.1%; Score 37; DB 16; Length 414;  
Best Local Similarity 70.0%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
Db 193 ILAGDHRQLP 202

RESULT 14  
US-10-437-963-166322  
Sequence 166322, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-22(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 166322  
LENGTH: 626  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)...(626)  
OTHER INFORMATION: unsure at all xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_65042C.1.pcp  
US-10-437-963-166322

Query Match 84.1%; Score 37; DB 16; Length 626;  
Best Local Similarity 70.0%; Pred. No. 16;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;



QY 3 ILAGDXXQLP 12  
:|||||  
Db 377 VLAGDHLQLP 386

## RESULT 15

US-10-369-493-2884  
; Sequence 2884, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2884  
; LENGTH: 650  
; TYPE: PRT  
; ORGANISM: Thermotoga maritima  
US-10-369-493-2884

Query March 84.1%; Score 37; DB 15; Length 650;  
Best Local Similarity 70.0%; Pred. No. 17;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
:|||||  
Db 404 VLAGDHLQLP 413

Search completed: April 18, 2005, 09:04:06  
Job time : 42.5693 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 07:32:05 ; Search time 11.562 Seconds  
(without alignments)  
99.861 Million cell updates/sec

Title: US-10-652-334-4  
Perfect score: 44  
Sequence: 1 XXILAGDXQLP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	86.4	635	2	T02699
2	38	86.4	642	2	D65085
3	38	86.4	648	2	C69423
4	38	86.4	656	2	E71080
5	38	86.4	989	2	T48845
6	38	86.4	993	2	A47500
7	38	86.4	993	2	G35633
8	37	84.1	650	2	G72429
9	37	84.1	653	2	B75105
10	37	84.1	683	2	S34700
11	37	84.1	751	2	C64367
12	36	81.8	530	2	D70476
13	36	81.8	663	2	H64312
14	35	79.5	774	2	D83208
15	34	77.3	360	2	G64575
16	34	77.3	542	2	T23869
17	34	77.3	602	2	T23866
18	34	77.3	633	2	T28788
19	34	77.3	640	2	S68546
20	34	77.3	692	2	E90113
21	34	77.3	693	2	T26415
22	34	77.3	1076	2	B96682
23	34	77.3	1289	2	B72354
24	33	75.0	191	2	T47242
25	33	75.0	274	2	D64588
26	33	75.0	472	2	A28701
27	33	75.0	505	2	JN0760
28	33	75.0	581	2	B81107
29	33	75.0	581	2	B81909

30	33	75.0	715	2	G75339	probable exodeoxyr
31	33	75.0	723	2	S46744	pif1 protein homol
32	33	75.0	744	2	A81719	exodeoxyribonuclea
33	33	75.0	746	2	D71564	probable exodeoxyr
34	33	75.0	798	2	A81263	exodeoxyribonuclea
35	33	75.0	798	2	AG1625	exodeoxyribonuclea
36	33	75.0	805	2	T40739	rm3-pif1 helicase
37	33	75.0	805	2	T47241	hypotheical prote
38	33	75.0	814	2	T00740	DNA helicase [limp
39	33	75.0	821	2	C84304	hypotheical prote
40	33	75.0	834	2	D68842	protein FR22.16 l
41	33	75.0	1075	2	C96682	hypotheical prote
42	33	75.0	1077	2	T50697	Dna2p - fission ye
43	33	75.0	1397	2	T51292	hypotheical helic
44	33	75.0	1398	2	T39568	hypotheical prote
45	33	75.0	1444	2	B84809	hypotheical prote

## ALIGNMENTS

## RESULT 1

T02699  
probable helicase At2g03270 [imported] - Arabidopsis thaliana

N:Alternate names: hypothetical protein T1812.6

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004

C:Accession: T02699; D84446

R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Ron

submitted to the EMBL Data Library, September 1998

A:Description: Arabidopsis thaliana chromosome II BAC T1812 genomic sequence.

A:Reference number: Z14702

A:Accession: T02699

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-635 <RD>

A:Cross-references: UNIPROT:O81047; EMBL:AC005313; NID:G3548797; PID:G3548803

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.

enus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; MUID:20083487; PMID:10617197

A:Accession: D84446

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-635 <STD>

A:Cross-references: GB:AB02093; NID:G4335770; PID:NAD17447.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g03270; T1812.6

A:Map position: 2

C:Superfamily: probable DNA helicase MJ0104

Query Match 86.4%; Score 38; DB 2; Length 635;

Best Local Similarity 80.0%; Pred. NO. 1.5;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXQLP 12  
Db 386 ILAGDXQLP 395

## RESULT 2

transcription control factor enhancer-binding protein - Methanobacterium thermoautotroph

C:Species: Methanobacterium thermoautotrophicum

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C:Accession: D69085

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

Qiu, D.; Spadator, R.; Viscare, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivani, N.

Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
 A:Reference number: A69000; MUID:98037514; PMID:9371463  
 A:Accession: D69085  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-642 <MTH>  
 A:Cross-references: UNIPROT:O27671; GB:AE000922; GB:AE000666; NID:g2622754; PIDN:AA8610  
 A:Experimental source: strain Delta H  
 C:Genetics:  
 A:Gene: MTH1634  
 C:Superfamily: probable DNA helicase MJ0104

Query Match 86.4%; Score 38; DB 2; Length 642;  
 Best Local Similarity 80.0%; Pred. No. 1.5;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ILAGDXXQLP 12  
 Db 390 ILAGDHRQLP 399

## RESULT 3

DNA helicase homolog - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
 C:Accession: C69423  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kyrness, E.F.  
 .; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 .; Nature 390, 364-370, 1997  
 A:Authors: Utecherback, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
 A:Reference number: A69250; MUID:98043343; PMID:9389475  
 A:Accession: C69423  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-648 <KLE>  
 A:Cross-references: UNIPROT:O28883; GB:AE001009; GB:AE000782; NID:g2689332; PIDN:AA86986  
 C:Superfamily: probable DNA helicase MJ0104

Query Match 86.4%; Score 38; DB 2; Length 648;  
 Best Local Similarity 80.0%; Pred. No. 1.5;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ILAGDXXQLP 12  
 Db 398 ILAGDHRQLP 407

## RESULT 4

probable DNA-binding protein - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 12-Jul-2004  
 C:Accession: E71080  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekit  
 M.; Onfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi  
 DNA Res. 5, 55-76, 1998  
 A>Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
 A:Reference number: A71000; MUID:98344137; PMID:9679194  
 A:Accession: E71080  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-656 <KAW>  
 A:Cross-references: UNIPROT:O58624; GB:AP000004; NID:g3236131; PIDN:BA30003.1; PID:g325  
 A:Experimental source: strain OT3  
 A>Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C:Genetics:  
 A:Gene: PH0909

Query Match 86.4%; Score 38; DB 2; Length 656;

Best Local Similarity 80.0%; Pred. No. 1.5;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ILAGDXXQLP 12  
 Db 405 ILAGDHRQLP 414

## RESULT 5

insulin II gene enhancer-binding protein rlp1 [validated] - golden hamster  
 C:Species: Mesocricetus auratus (golden hamster)  
 C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
 C:Accession: T48845  
 R:Shieh, S.Y.; Stellrecht, C.M.; Teal, M.J.  
 J. Biol. Chem. 270, 21503-21508, 1995  
 A>Title: Molecular characterization of the rat insulin enhancer-binding complex 3b2. Clor  
 A:Reference number: 224545; MUID:95394901; PMID:7665561  
 A:Accession: T48845  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-989 <SHI>  
 A:Cross-references: UNIPROT:Q60560; EMBL:L15625; NID:g290918; PIDN:AA800104.1; PID:g2909;  
 A:Experimental source: tissue type insulinoma  
 C:Genetics:  
 A>Note: Rlp1  
 C:Complex: rlp1 is a component of the Rlpe3b-binding complex 3b2, which is composed of at  
 A:Description: the 3b2 complex is one of two complexes, binding to the Rlpe3b element of  
 ent Rlpe3a [validated, MUID:95394901]

Query Match 86.4%; Score 38; DB 2; Length 989;  
 Best Local Similarity 80.0%; Pred. No. 2.4;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ILAGDXXQLP 12  
 Db 395 ILAGDHRQLP 404

## RESULT 6

Ig mu chain switch region binding protein 2 - human  
 A:Accession: A47500  
 A:Title: The human SmuBP-2, a DNA-binding protein specific to the single-stranded guanine  
 A:Reference number: A47500; MUID:93352537; PMID:8349627  
 A:Accession: A47500  
 A:Molecule type: mRNA  
 A:Residues: 1-993 <FKU>  
 A:Cross-references: UNIPROT:P38935; GB:L14754  
 R:Kerr, D.; Khalili, K.  
 J. Biol. Chem. 268, 15876-15881, 1993  
 A>Title: A recombinant cDNA derived from human brain encodes a DNA binding protein that  
 A:Reference number: A40804; MUID:91340730; PMID:1174899

A:Accession: A40804  
 A:Molecule type: mRNA  
 A:Residues: 'GGRV', 495-862, 'K', 864-865, 'T' <KER>  
 A:Cross-references: GB:M64979; NID:g183249; PIDN:AA58611.1; PID:g183250  
 A:Experimental source: brain stem  
 A>Note: the authors translated the codon CAG for residue 507 as Glu  
 C:Genetics:  
 A:Gene: GDB:IGMBP2  
 A:Cross-references: GDB:434023; OMIM:600502  
 A:Map position: 11q13.2-11q13.4  
 C:Keywords: immunoglobulin, single-stranded DNA binding

Query Match 86.4%; Score 38; DB 2; Length 993;  
 Best Local Similarity 80.0%; Pred. No. 2.4;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
:|||||  
Db 396 ILAGDHRQLP 405

## RESULT 7

S35633  
DNA-binding protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S35633  
R:Mutsu, T.R.; Fukita, Y.; Miyoshi, T.; Shimizu, A.; Honjo, T.  
Nucleic Acids Res. 21, 1761-1766, 1993  
A:Title: Isolation of cDNA encoding a binding protein specific to 5'-phosphorylated sing  
A:Reference number: S35633; MUID:93261806; PMID:8493094  
A:Accession: S35633  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-993 <MTZ>  
A:Cross-references: UNIPROT:P40694; GB:L10075; NID:g293805; PIDN:AAA40143.1; PID:g293806  
C:Keywords: DNA binding

Query Match 86.4%; Score 38; DB 2; Length 993;  
Best Local Similarity 80.0%; Pred. No. 2.4;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
:|||||  
Db 395 ILAGDHRQLP 404

## RESULT 8

G72429  
hypothetical protein TM0005 - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: G72429  
R:Neilson, K.B.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey  
Garrett, M.M.; Stewart, A.M.; Colton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.J.  
C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
A:Reference number: A72200; MUID:99287316; PMID:10360571  
A:Accession: G72429  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-650 <ARN>  
A:Cross-references: UNIPROT:Q9WXM0; GB:AE001689; GB:AE000512; NID:g4980483; PIDN:AA03505  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TM0005  
C:Superfamily: probable DNA helicase MJ0104

Query Match 84.1%; Score 37; DB 2; Length 650;  
Best Local Similarity 70.0%; Pred. No. 2.6;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
:|||||  
Db 404 VLADGHRQLP 413

## RESULT 9

B75105  
probable DNA helicase PAB1561 - Pyrococcus abyssi (strain Orsay)  
C:Species: Pyrococcus abyssi  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: B75105  
R:anonymou, Genoscope  
submitted to the EMBL Data Library, July 1999  
A:Description: Pyrococcus abyssi genome sequence: Insights into archaeal chromosome stru

A:Reference number: A75001  
A:Accession: B75105  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-653 <KAN>  
A:Cross-references: UNIPROT:Q9UZB6; GB:AJ248286; GB:AL096836; NID:g5458366; PIDN:CAB5014  
A:Experimental source: strain Orsay  
C:Genetics:  
A:Gene: PAB1561  
C:Superfamily: probable DNA helicase MJ0104

Query Match 84.1%; Score 37; DB 2; Length 653;  
Best Local Similarity 70.0%; Pred. No. 2.6;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
:|||||  
Db 405 VLADGHRQLP 414

## RESULT 10

S34700  
probable purine nucleotide-binding protein YKL017c - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: S34700; S37830; S37834  
R:Wiemann, S.; Voss, H.; Schwager, C.; Rupp, T.; Stegemann, J.; Zimmermann, J.; Grothues  
submitted to the EMBL Data Library, July 1993  
A:Description: Sequencing and analysis of 51.5 kilobases on the left arm of chromosome X  
A:Reference number: S34679  
A:Accession: S34700  
A:Molecule type: DNA  
A:Residues: 1-683 <MTB>  
A:Cross-references: UNIPROT:P34243; EMBL:X74152; NID:g450363; PID:g395256  
A:Experimental source: strain S288C  
R:Wiemann, S.; Voss, H.; Schwager, C.; Rupp, T.; Grothues, D.; Sensen, C.; Stegemann, J.  
submitted to the Protein Sequence Database, March 1994  
A:Reference number: S37825  
A:Accession: S37830  
A:Molecule type: DNA  
A:Residues: 1-683 <W12>  
A:Cross-references: EMBL:Z28017; NID:g486006; PID:g486007; MIPS:YKL017C  
A:Experimental source: strain S288C  
R:Rieger, M.  
submitted to the Protein Sequence Database, March 1994  
A:Reference number: S37832  
A:Accession: S37834  
A:Molecule type: DNA  
A:Residues: 1-683 <RTB>  
A:Cross-references: EMBL:Z28017; NID:g486006; PID:g486007; MIPS:YKL017C  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:DIP1  
A:Cross-references: SGD:S0001500; MIPS:YKL017C  
A:Map position: 11L  
C:Superfamily: probable DNA helicase MJ0104

C:Keywords: ATP; P-loop; purine nucleotide binding  
P123-236/Region: nucleotide-binding motif A (P-loop)  
P123/Binding site: ATP/GTP (Lys) #status predicted

Query Match 84.1%; Score 37; DB 2; Length 683;  
Best Local Similarity 70.0%; Pred. No. 2.7;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 ILAGDXXQLP 12  
:|||||  
Db 420 VLADGHRQLP 429

## RESULT 11

C84367  
DNA binding protein eukaryotic-like [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: C84367  
R/Seq: W.V.; Kennedy, S.P.; Mahliras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leitner, B.; Keller, K.; Cruz, R.; Danon, M.J.; Hough, D.W.; Madocks, D.G.; Jbdic  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; LA  
A/Title: Genome sequence of Halobacterium species NRC-1.  
A/Reference number: A84160; MUID:20504483; PMID:11016950  
A/Accession: C84367  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-751 <STO>  
A/Cross-references: UNIPROT:Q9HNB7; GB:AE004437; NID:g10581584; PIDN:AAG20303.1; GSPDB:C  
C/Genetics:  
A/Gene: dbp

Query Match 84.1%; Score 37; DB 2; Length 751;  
Best Local Similarity 70.0%; Pred. No. 3;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

3 ILAGDXXQLP 12  
Db 501 VLADHRLP 510

RESULT 12  
D70476  
DNA helicase - Aquifex aeolicus  
C/Species: Aquifex aeolicus  
C/Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C/Accession: D70476  
R/Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O  
V.  
Nature 392, 353-358, 1998  
A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A/Reference number: A70300; MUID:98196666; PMID:9537320  
A/Accession: D70476  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-530 <ADP>  
A/Cross-references: UNIPROT:O67840; GB:AE000770; NID:g2984274; PIDN:AAC07803.1; PID:g298  
A/Experimental source: strain VFS  
C/Genetics:  
A/Gene: helix

Query Match 81.8%; Score 36; DB 2; Length 530;  
Best Local Similarity 70.0%; Pred. No. 3.5;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

3 ILAGDXXQLP 12  
Db 283 IMAGDHQLP 292

RESULT 13  
H64312  
probable DNA helicase M0104 - Methanococcus jannaschii  
C/Species: Methanococcus jannaschii  
C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C/Accession: H64312  
R/Balt, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
J.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
rsen, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurest, M.A.  
Science 273, 1058-1073, 1996  
A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A/Reference number: A64300; MUID:96337999; PMID:8688087  
A/Accession: H64312  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-663 <BDU>  
A/Cross-references: UNIPROT:O57568; GB:U67467; GB:L77117; NID:g2826242; PIDN:AAB98084.1;

C/Genetics:  
A/Map position: FOR99243-101234  
A/Start codon: TTG  
C/Superfamily: probable DNA helicase M0104

Query Match 81.8%; Score 36; DB 2; Length 663;  
Best Local Similarity 70.0%; Pred. No. 4.5;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

3 ILAGDXXQLP 12  
Db 422 IMAGDHQLP 431

RESULT 14  
D83208  
probable ferredoxin PA3491 [imported] - Pseudomonas aeruginosa (strain PA01)  
C/Species: Pseudomonas aeruginosa  
C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C/Accession: D83208  
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A/Reference number: A82950; MUID:20437337; PMID:10984043  
A/Accession: D83208  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-774 <STO>  
A/Cross-references: UNIPROT:Q9HYB8; GB:AE004770; GB:AE004091; NID:g9949633; PIDN:AAG06875  
A/Experimental source: strain PA01  
C/Genetics:  
A/Gene: PA3491  
C/Superfamily: probable iron-sulfur binding NADH dehydrogenase YPO2244; ferredoxin 2[Fe-

Query Match 79.5%; Score 35; DB 2; Length 774;  
Best Local Similarity 60.0%; Pred. No. 9.1;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

3 ILAGDXXQLP 12  
Db 352 ILAGDXXQLP 361

RESULT 15  
G64575  
conserved hypothetical protein HP0447 - Helicobacter pylori (strain 26695)  
C/Species: Helicobacter pylori  
C/Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C/Accession: G64575  
R/Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKenney  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney, L.  
Nature 388, 539-547, 1997  
A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C.  
A/Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A/Reference number: A64520; MUID:97394467; PMID:9252185  
A/Accession: G64575  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-360 <TOM>  
A/Cross-references: UNIPROT:O25195; GB:AE000560; GB:AE000511; NID:g2313554; PIDN:AAD0751.  
C/Genetics:  
A/Start codon: GTG

Query Match 77.3%; Score 34; DB 2; Length 360;  
Best Local Similarity 70.0%; Pred. No. 6.6;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

3 ILAGDXXQLP 12  
Db 96 ILVGDHRLP 105

Tue Apr 19 09:10:42 2005

us-10-652-334-4.rpr

Page 5

Search completed: April 18, 2005, 08:06:01  
Job time : 12.562 secs

---

This page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 07:33:55 ; Search time 50.1022 Seconds

(without alignments)  
122.648 Million cell updates/sec

Title: US-10-652-334-4

Perfect score: 44

Sequence: 1 XXILAGDXQLP 12

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Uniprot\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	88.6	471	2	Q7MV14 porphyromon
2	38	86.4	634	2	Q72RC7 leptospira
3	38	86.4	634	2	Q6F4E3 leptospira
4	38	86.4	635	2	Q6F4E3 leptospira
5	38	86.4	639	2	Q6F4E3 leptospira
6	38	86.4	642	2	Q27671 methanobact
7	38	86.4	648	2	Q28883 archaeoglob
8	38	86.4	655	2	Q8U398 pyrococcus
9	38	86.4	656	2	Q58624 pyrococcus
10	38	86.4	695	2	Q6FKP9 candida gla
11	38	86.4	988	1	Q9EQN5 ratiu nov
12	38	86.4	989	1	Q60560 mesocricetu
13	38	86.4	993	1	SMB2_HUMAN
14	38	86.4	993	1	SMB2_MOUSE
15	38	86.4	1190	2	Q7S8B9 mus musculu
16	37	84.1	650	2	Q9WXM0 thermocoga
17	37	84.1	651	2	Q6F2U8 oryza sativ
18	37	84.1	652	2	Q8A613 bacteroides
19	37	84.1	653	2	Q9UZB6 pyrococcus
20	37	84.1	683	1	YKB7_YEAST
21	37	84.1	716	2	Q7S1L9 bacteriophag
22	37	84.1	751	2	Q9HNB7 halobacteri
23	37	84.1	763	2	Q7WMP1 rhodospirill
24	36	81.8	413	2	Q7QY97 gladiolus lam
25	36	81.8	530	2	Q67840 aquifex aeo
26	36	81.8	633	2	Q6LWX3 methanococ
27	36	81.8	649	2	Q6CFH6 yarrowia li
28	36	81.8	663	1	Y104_METTA
29	36	81.8	681	2	Q6CMT3 kluyveromyc
30	36	81.8	698	2	Q8T269 methanopyru
31	36	81.8	734	2	Q7QYL6 giardia lam

32	36	81.8	1024	2	Q75JAO dictyosteli
33	35	79.5	530	2	Q9QSK3 chilo iride
34	35	79.5	570	2	Q6FAS6 actinobact
35	35	79.5	640	1	PPCM_MOUSE
36	35	79.5	657	2	Q7SCL7 aethya goos
37	35	79.5	755	2	Q6BK27 debaryomyce
38	35	79.5	765	2	Q8RCB3 thermotaneer
39	35	79.5	774	1	RNFC_PSEBAE
40	34	77.3	50	2	Q8EAV3 pseudomonas
41	34	77.3	242	2	Q6N3J4 shewanella
42	34	77.3	360	2	Q2S195 rhodospseudo
43	34	77.3	457	2	Q2S195 helicobacte
44	34	77.3	472	2	Q86U01 homo sapien
45	34	77.3	542	2	Q8G3N4 bifidobacte
					Q09594 caenorhabdi

## ALIGNMENTS

RESULT 1					
ID	Q7MV14	PRELIMINARY;	PRT;	471 AA.	
AC	Q7MV14				
DT	01-MAR-2004 (TREMBLrel. 26, Created)				
DT	01-MAR-2004 (TREMBLrel. 26, Last sequence update)				
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)				
DE	Helicase, putative.				
CN	OrderedLocustNames=PG1303;				
OS	Porphyromonas gingivalis (Bacteroides gingivalis).				
OC	Bacteriia; Bacteroidetes; Bacteroides (class); Bacteroidales;				
OC	Porphyromonadaceae; Porphyromonae.				
OK	NCBI_TaxID=837;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=W83;				
RX	MEDLINE=22829867; PubMed=12949112;				
RX	DOI=10.1128/JB.185.18.5591-5601.2003;				
RA	Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.B.,				
RA	Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,				
RA	Haft D.H., Kolonay J.P., Nelson W.C., Mason T.M., Tallon L., Gray J.,				
RA	DeAngelis D., Tettelin H., Dong H., Galvin J.L., Duncan M.C.,				
RA	Devlin F.B., Fraser C.M.;				
RT	"Complete genome sequence of the oral pathogenic bacterium				
RT	Porphyromonas gingivalis strain W83."				
RL	J. Bacteriol. 185:5591-5601(2003).				
DR	EMBL; AB017176; AAC6376.1; -.				
DR	TIGR; PG1303; -.				
DR	GO; GO:0004386; F:helicase activity; IEA.				
KW	Complete proteome; Helicase.				
SQ	SEQUENCE 471 AA; 52502 MW; FA887B1BB829D761 CRC64;				
Query Match 88.6%; Score 39; DB 2; Length 471;					
Best Local Similarity 80.0%; Pred. No. 2.6;					
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;					
QY	3 ILAGDXQLP 12				
DB	163 ILAGDMAQLP 172				
RESULT 2					
ID	Q72RC7	PRELIMINARY;	PRT;	634 AA.	
AC	Q72RC7				
DT	05-JUL-2004 (TREMBLrel. 27, Created)				
DT	05-JUL-2004 (TREMBLrel. 27, Last sequence update)				
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)				
DE	Hypothetical protein.				
CN	OrderedLocustNames=LIC11819;				
OS	Leptospira interrogans (serogroup Icterohaemorrhagiae / serovar				
OS	Copenhagen).				
OC	Bacteriia; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.				
OK	NCBI_TaxID=44275;				

[1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Piocruz LI-130;  
 RX PubMed=15028702; DOI=10.1128/JB.186.7.2164-2172.2004;  
 RA Nacimento A.L.T.O., Ko A.I., Martins E.A.L., Monteiro-Vitorello C.B.,  
 RA Ho P.L., Haake D.A., Verjovski-Almeida S., Hartskeerl R.A.,  
 RA Marques M.V., Oliveira M.C., Menck C.F.M., Leite L.C.C., Carver H.,  
 RA Coutinho L.L., Degreve W.M., Dellagostin O.A., El-Dorry H.,  
 RA Ferro E.S., Ferro M.I.T., Furian L.R., Gambellini M., Gigliotti E.A.,  
 RA Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R.,  
 RA Jeronimo S.M.B., Junqueira-de-Azevedo I.L.M., Kimura E.T.,  
 RA Kuramae E.E., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R.,  
 RA de Oliveira R.C., Pereira G.G., Reis M.S., Schrieffer A.,  
 RA Siqueira W.J., Sommer P., Tsai S.M., Simpson A.J.G., Ferro J.A.,  
 RA Camargo L.E.A., Kiteajima J.P., Setubal J.C., Van Sluys M.A.;  
 RT "Comparative genomics of two Leptospira interrogans serovars reveals  
 RT novel insights into physiology and pathogenesis.";  
 RL J. Bacteriol. 186:2164-2172(2004).  
 DR EMBL; AE017294; AA570407.1; -  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR000212; VWD-helicase.  
 DR Pfam; PF00580; VWD-helicase; 1.  
 DR SMART; SM00382; AAA; 1.  
 KM ATP-binding; Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 634 AA; 7256 MW; 8628954412AF267 CRC64;

Query Match 86.4%; Score 38; DB 2; Length 634;  
 Best Local Similarity 80.0%; Pred. No. 6.3;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ILAGDXXQLP 12  
 Db 395 ILAGDHHQLP 404

RESULT 3  
 ID 08FAE3 PRELIMINARY; PRT; 634 AA.  
 AC 08FAE3;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 23 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Putative helicase.  
 GN OrderedLocustNames=LA2098;  
 OS Leptospira interrogans.  
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
 NCBI\_TaxId=173;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar 1a1;  
 RX MEDLINE=22596143; PubMed=12712204; DOI=10.1038/nature01597;  
 RA Ren S.-X., Pu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,  
 RA Zhang Y.-X., Xiong H., Lu L.-F., Jiang H.-O., Jia J., Tu Y.-F.,  
 RA Jiang J.-X., Gu W.-Y., Zhang Y.-O., Cai Z., Sheng H.-H., Yin H.-F.,  
 RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,  
 Yeo Z.-U., Shen G., Qiang B.-O., Xia Q.-C., Guo X.-K., Danchin A.,  
 RA Saint Glrons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,  
 RA Xu J.-G., Zhao G.-P.;  
 RT "Unique physiological and pathogenic features of Leptospira  
 RT interrogans revealed by whole-genome sequencing.";  
 RL Nature 432:888-893(2003).  
 DR EMBL; AE011381; AA49297.1; -  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004003; F:ATP-dependent DNA helicase activity; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR GO; GO:0006281; P:DNA repair; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR011545; DEAD/DEAH\_N.  
 DR InterPro; IPR000212; VWD-helicase.  
 DR Pfam; PF00580; VWD-helicase; 1.

DR SMART; SM00382; AAA; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 KM ATP-binding; Complete proteome.  
 SQ SEQUENCE 634 AA; 7260 MW; 4A6928E206F93590 CRC64;

Query Match 86.4%; Score 38; DB 2; Length 634;  
 Best Local Similarity 80.0%; Pred. No. 6.3;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ILAGDXXQLP 12  
 Db 395 ILAGDHHQLP 404

RESULT 4  
 ID 09FNX9 PRELIMINARY; PRT; 635 AA.  
 AC 09FNX9;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Putative helicase.  
 GN Name=atpc-2 gene;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxId=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole plant;  
 RA Bezhani S., Sheremeti I., Pfannschmidt T., Oelmüller R.;  
 RT "A repressor with similarities to pro- and eukaryotic DNA helicases  
 RT controls the assembly of the CAP-box binding complex at a  
 RT photosynthesis gene promoter.";  
 RL J. Biol. Chem. 10:1074-1074(2001).  
 DR EMBL; AJ300306; CAC16347.1; -  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0004386; F:helicase activity; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR011545; DEAD/DEAH\_N.  
 DR InterPro; IPR004483; put\_DNA\_helic.  
 DR SMART; SM00382; AAA; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR TIGRFAMs; TIGR00376; put\_DNA\_helic; 1.  
 KM ATP-binding; Helicase.  
 SQ SEQUENCE 635 AA; 70335 MW; 6149646686631P29 CRC64;

Query Match 86.4%; Score 38; DB 2; Length 635;  
 Best Local Similarity 80.0%; Pred. No. 6.3;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ILAGDXXQLP 12  
 Db 386 ILAGDHHQLP 395

RESULT 5  
 ID 081047 PRELIMINARY; PRT; 639 AA.  
 AC 081047; 094AT0;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE Putative helicase.  
 GN Name=At2G03270;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxId=3702;

```

RN [1]
RN SEQUENCE FROM N.A.
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., Vanaken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carreir A.J., Greasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RA Ronney S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA Shenn M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamuya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miyata M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shim P., Southwick A., Shinzaki K., Davys R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamuya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyer M.C., Mizuda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shim P., Southwick A., Shinzaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AC006284; AAD17447.2; -
DR EMBL: AC005313; AAM15033.1; -
DR EMBL: AY045820; AAK76494.1; -
DR EMBL: AY091361; AAM14300.1; -
DR PIR: T02699; T02699.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0004386; F:helicase activity; IEA.
DR GO: GO:000166; F:nucleotide binding; IEA.
DR InterPro: IPR003593; AAA_Atpase.
DR InterPro: IPR011545; DEAD/DEAH N.
DR InterPro: IPR004483; put_DNA_helic.
DR SMART: SM00382; AAA; 1.
DR SMART: SM00487; DEXDC; 1.
DR TIGRfams: TIGR00376; put_DNA_helic; 1.
DR ATP-binding; Helicase.
SQ SEQUENCE 639 AA; 70850 MW; C91340DAE84C978C CRC64;

```

```

OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RN SEQUENCE FROM N.A.
RA STRAIN=Delta H;
RC MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pochter B., Qiu D.,
RA Spadator R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McQuigall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mo J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155 (1997).
DR EMBL: AB000922; AAB86107.1; -
DR PIR: D69085; D69085.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:000166; F:nucleotide binding; IEA.
DR InterPro: IPR003593; AAA_Atpase.
DR InterPro: IPR004483; put_DNA_helic.
DR SMART: SM00382; AAA; 1.
DR TIGRfams: TIGR00376; put_DNA_helic; 1.
DR ATP-binding; Complete proteome.
KV SEQUENCE 642 AA; 72788 MW; 278DA65037229377 CRC64;

```

Query Match 86.4%; Score 38; DB 2; Length 642;  
 Best Local Similarity 80.0%; Pred. No. 6.3;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy 3 ILAGDXXQLP 12
Db 390 ILAGDHHQLP 399

```

```

RESULT 7
ID 028883; PRELIMINARY; PRT; 648 AA.
AC 028883;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE DNA helicase, putative.
GN OrderedLocusNames=AF1388;
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RC MEDLINE=98049343; PubMed=9384975; DOI=10.1038/37052;
RA Klein H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Kerlbaum K.A., Dodson R.J., Gwinn M.L., Hickey E.K., Peterson J.D.,
RA Fleischmann R.D., Kierlavage A.R., Graham D.S., Kyprides N.C.,
RA Kirness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.J.,
RA Peterson S.N., Reich C.I., McNeil L.K., Badger J.H., Glodex A.,
RA Zhou L., Overbeek R., Gocayne J.D., Weidman J.F., McDonald L.A.,
RA Uterback T.R., Cotton M.D., Spriggs T., Artlich P., Kaine B.P.,
RA Sykes S.M., Sadow P.W., D'Andrea K.P., Bowman C., Fujii C.,
RA Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O.,
RA Woese C.R., Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370 (1997).
DR EMBL: AB001009; AAB89860.1; -
DR PIR: C69423; C69423.
DR TIGR: AF1388; -

```

DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR GO: GO:0004386; F:helicase activity; IEA.  
 DR GO: GO:0000166; F:nucleotide binding; IEA.  
 DR InterPro: IPR003593; AAA\_Arpase.  
 DR InterPro: IPR004483; put\_DNA\_helic.  
 DR SMART: SM00382; AAA; 1.  
 DR TIGRFAMs: TIGR00376; put\_DNA\_helic; 1.  
 KW ATP-binding: Complete proteome; Helicase.  
 SQ SEQUENCE 648 AA; 73775 MW; B507450785647D5 CRC64;

Query Match 86.4%; Score 38; DB 2; Length 648;  
 Best Local Similarity 80.0%; Pred. No. 6.4;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ILADXXQLP 12  
 Db 398 ILADHRLP 407

RESULT 8  
 ID Q8U398 PRELIMINARY; PRT; 655 AA.  
 AC Q8U398;  
 DT 01-JUN-2002 (TREMblrel. 21, Created)  
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)  
 DE Dna2-nam7 helicase family protein.  
 OS OrderedLocustNames=PR0572;  
 GN Pyrococcus furiosus.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcaceae;  
 OC Pyrococcus.  
 NCBI\_TaxID=2261;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;  
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;  
 RT "The complete sequence of the Pyrococcus furiosus genome."  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB010180; F:AB0696.1;  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR SMART: SM00487; DEXDC; 1.  
 DR TIGRFAMs: TIGR00376; put\_DNA\_helic; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 655 AA; 74581 MW; 0AB579DA14ABEBD0 CRC64;

Query Match 86.4%; Score 38; DB 2; Length 655;  
 Best Local Similarity 80.0%; Pred. No. 6.5;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ILADXXQLP 12  
 Db 405 ILADHRLP 414

RESULT 9  
 ID OS6624 PRELIMINARY; PRT; 656 AA.  
 AC OS6624;  
 DT 01-AUG-1998 (TREMblrel. 07, Created)  
 DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)  
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)  
 DE 656aa long hypothetical DNA-binding protein.  
 GN OrderedLocustNames=PH0909;  
 OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcaceae;  
 OC Pyrococcus.  
 NCBI\_TaxID=53953;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3;  
 RX MEDLINE=98344137; PubMed=9679194;

RA Kawarabayashi Y., Sawada M., Horikawa H., Hatake Y., Hino Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Koeugi H., Hosoyama A., Nagai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamita M., Ofuku Y.,  
 RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kushta N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 RA Masuchi Y., Shizuya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-  
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3."  
 RL DNA Res. 5:55-76 (1998).  
 DR EMBL: AP000004; BAA3003.1;  
 DR PIR: E71080; E71080.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR InterPro: IPR004483; put\_DNA\_helic.  
 DR TIGRFAMs: TIGR00376; put\_DNA\_helic; 1.  
 KW Complete proteome; DNA-binding; Hypothetical protein.  
 SQ SEQUENCE 656 AA; 74922 MW; 252C46BFF3BBAPOF CRC64;

Query Match 86.4%; Score 38; DB 2; Length 656;  
 Best Local Similarity 80.0%; Pred. No. 6.5;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ILADXXQLP 12  
 Db 405 ILADHRLP 414

RESULT 10  
 ID Q6FKP9 PRELIMINARY; PRT; 695 AA.  
 AC Q6FKP9;  
 DT 05-JUL-2004 (TREMblrel. 27, Created)  
 DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMblrel. 27, Last annotation update)  
 DE Similar to sp|P34243|Saccharomyces cerevisiae YK017c DIPI.  
 GN ORFNames=CAGJ0109735g;  
 OS Candida glabrata CBS138.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 NCBI\_TaxID=284593;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS138;  
 RA Genolevures;  
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,  
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barde V.,  
 RA Batray S., Blanchin S., Beckerich J.M., Beyne B., Bleykasten C.,  
 RA Boissarme A., Boyer J., Cattolico L., Confiantier F., de Daruvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppe A.,  
 RA Hantre F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrist A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicaud J.M., Nikolaki M., Oztas S., Ozler-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straud M.L., Suleau A.,  
 RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof B., Wirth B.,  
 RA Zenlou-Meyer M., Zivanovic I., Bolocin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,  
 RA Wincker P., Soulet J.L.;  
 RT "Genome evolution in yeasts."  
 RL Nature 430:35-44 (2004).  
 DR EMBL: CR380958; CAG62165.1;  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR InterPro: IPR011545; DEAD/DEAH N.  
 DR InterPro: IPR004483; put\_DNA\_helic.  
 DR SMART: SM00487; DEXDC; 1.  
 DR TIGRFAMs: TIGR00376; put\_DNA\_helic; 1.  
 SQ SEQUENCE 695 AA; 78391 MW; DC4995FE4CDOB420 CRC64;

Query Match 86.4%; Score 38; DB 2; Length 695;  
 Best Local Similarity 70.0%; Pred. No. 6.9;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy 3 ILAGDXXQLP 12
    :|||||
Db 425 VLADGSKQLP 434

RESULT 11
Q9EON5 PRELIMINARY; PRT; 988 AA.
ID 09EON5
AC 09EON5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
OS Antifreeze-enhancer binding protein ABP.
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OC NCB1_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=20558263; PubMed=11106437;
RA Miao M., Chan S.L., Fletcher G.L., Hew C.L.;
RT "The rat ortholog of the presumptive flounder antifreeze enhancer-
RL binding protein is a helicase domain-containing protein.";
DR EMBL; AF199411; AAC28561.1; -.
DR HSSP; P38935; IMSZ.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:000166; P:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_Arase.
DR InterPro; IPR011545; DEAD/DEAF_N.
DR InterPro; IPR004483; put_DNA_helic.
DR InterPro; IPR001374; R3H.
DR InterPro; IPR000058; Znt_AN1.
DR Pfam; PF01424; R3H; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00393; R3H; 1.
DR SMART; SM00154; ZNF_AN1; 1.
DR TIGRFAMs; TIGR00376; put_DNA_helic; 1.
DR ATP-binding.
SQ SEQUENCE 988 AA; 108439 MW; 624BD6122F8C8D8A CRC64;

Query Match 86.4%; Score 38; DB 2; Length 988;
Best Local Similarity 80.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
RA Shieh S.-Y., Stellierecht C.M.M., Tsai M.-J.;
RT "Molecular characterization of the rat insulin enhancer-binding
RT complex 3b2. Cloning of a binding factor with putative helicase
RT motifs.";
RL J. Biol. Chem. 270:21503-21508(1995).
CC -1- FUNCTION: Binds to the insulin II gene R1PE3B enhancer region.
CC -1- SUBUNIT: Composed of at least three polypeptides: p58, p62, and
CC p110.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- TISSUE SPECIFICITY: High expression in brain and testis, moderate
CC in heart, spleen, and kidney, and low in other tissues.
CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.
CC -----
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
```

```
Qy 3 ILAGDXXQLP 12
    :|||||
Db 395 ILAGDHRQLP 404

RESULT 12
SMB2_MESAU STANDARD; PRT; 989 AA.
ID 060560;
AC 060560;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE DNA-binding protein SMUBP-2 (Immunoglobulin mu binding protein 2)
DE (SMUBP-2) (Insulin II gene enhancer-binding protein) (R1PE3B-binding
DE complex 3b2 p110 subunit) (R1P-1)
GN Name=IGHMBP2; Synonym=IRIP1; SMUBP2;
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OC NCB1_TaxId=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Insulinoma;
RX MEDLINE=95394901; PubMed=7665561; DOI=10.1074/jbc.270.37.21503;
```

```
RESULT 13
SMB2_HUMAN STANDARD; PRT; 993 AA.
ID P38935; Q14177;
AC P38935; Q14177;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE DNA-binding protein SMUBP-2 (Immunoglobulin mu binding protein 2)
DE (SMUBP-2) (Glial factor-1) (GF-1)
GN Name=IGHMBP2; Synonym=SMBP2; SMUBP2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCB1_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93352537; PubMed=8349627;
RA Fukita Y., Mizuta T.-R., Shirozu M., Ozawa K., Shimizu A., Honjo T.;
RT "The human smu bp-2, a DNA-binding protein specific to the single-
RT stranded guanine-rich sequence related to the immunoglobulin mu chain
RT switch region.";
RL J. Biol. Chem. 268:17463-17470(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99160766; PubMed=10049831; DOI=10.1006/viro.1998.9588;
```

Query Match	86.4%	Score 39,	DB 1,	Length 993,
Best Local Similarity	80.0%	Pred. No. 10,		
Matches	8;	Conservative	0;	Mismatches
			2;	Indels
			0;	Gaps
			0;	

```

Oy 3 ILAGDXXQLP 12
    |||||
    ||||
Db 396 ILAGDHQLP 405

RESULT 14
SMB2_MOUSE
ID SMB2_MOUSE STANDARD; PRT; 993 AA.
AC PA0654;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE DNA-binding protein SWBP-2 (Immunoglobulin M binding protein 2)
DS (SWBP-2) (Cardiac transcription factor 1) (CATF1).
GN Name=Ighmbp2; Synonyms=Smbp-2, Smbp2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Spleen;
RC MEDLINE=93261806; PubMed=8493094;
RA Maita T.-R., Fukita Y., Miyoshi T., Shimizu A., Honjo T.;
RT "Isolation of cDNA encoding a binding protein specific to 5'-
RT phosphorylated single-stranded DNA with G-rich sequences.";
RL Nucleic Acids Res. 21:1761-1766(1993).
CC -1- FUNCTION: DNA-binding protein specific to 5'-phosphorylated
CC single-stranded guanine-rich sequence related to the
CC immunoglobulin mu chain switch region. Preferentially binds to the
CC motif (5'-GGGCT-3').
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- TISSUE SPECIFICITY: In all tissues examined.
CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L10075; AAA40143.1; -.
DR PIR; S35633; S35633.
DR HSSP; P38935; IMSZ.
DR MGD; MGI:99954; Ighmbp2.
DR InterPro; IPR004483; putc_DNA_helic.
DR InterPro; IPR01374; R3H.
DR Pfam; PF01424; R3H; 1.
DR TIGRFAMs; TIGR00376; putc_DNA_helic; 1.
KW Activator; ATP-binding; DNA-binding; Helicase; Nuclear protein;
KW Transcription regulation.
KM NP_BIND 213 220 ATP (By similarity).
FT FT 249 425 Leu-rich.
FT FT 370 373 Poly-Val.
FT FT 637 783 SS DNA-binding (By similarity).
FT FT 793 861 Gln/Pro-rich.
FT FT 862 866 Poly-Lys.
FT FT 862 866 Nuclear localization signal (Potential).
SQ SEQUENCE 993 AA; 109466 MW; 2FA0850DBADE35B CRC64;

Query March 86.4%; Score 38; DB 1; Length 993;
Best Local Similarity 80.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 3 ILAGDXXQLP 12
    |||||
    ||||
Db 395 ILAGDHQLP 404

```

07S8B9  
 ID 07S8B9 PRELIMINARY; PRT: 1190 AA.  
 AC 07S8B9;  
 DT 01-MAR-2004 (TRENBLREL. 26, Created)  
 DT 01-MAR-2004 (TRENBLREL. 26, Last sequence update)  
 DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)  
 DE Hypothetical protein.  
 GN Name=NCU06627.1;  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OR74A;  
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
 Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,  
 Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,  
 RA Seitzmunkoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,  
 Koche G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,  
 Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,  
 RA Kamai M., Kanwysellis W., Mauceli E., Bielke C., Rudd S., Friseman D.,  
 RA Krystofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,  
 RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmant S.A.,  
 Desguza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,  
 RA Yarden O., Pliemann M., Seiler S., Dunlap J., Radford A., Aramayo R.,  
 RA Natvig D.O., Alex L.A., Mannheim G., Ebbole D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,  
 RA "The Genome Sequence of the Filamentous Fungus Neurospora crassa."  
 RL Nature 0:0-0(2003).  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AABX01000237; EAA32579.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1190 AA; 132937 MW; E3C5DP5A853E20A3 CRC64;

Query Match 86.4%; Score 38; DB 2; Length 1190;  
 Best Local Similarity 80.0%; Pred. No. 12;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Oy 3 ILAGDXQLP 12  
 |||||  
 Db 818 ILAGDVQLP 827

Search completed: April 18, 2005, 08:15:43  
 Job time : 52.1022 secs

**This Page Blank (uspto)**





CC Muscular Dystrophy etc. It can be used to identify disease conditions  
 CC involving a defect in the complex, by transfecting cells with encoding  
 CC nucleic acid and determining the proportion of defective complex before  
 CC and after transfection. It is also useful to screen for drugs involved in  
 CC peptidyl transferase activity during translation, inhibiting the  
 CC interaction between MT1 and eRF3 or involved in enhancing translation  
 CC termination. Vectors comprising polynucleotides encoding the complex (or  
 CC antisense sequences) can be constructed and introduced into cells to  
 CC interfere with complex expression and so modulate the efficiency of  
 CC translation termination of mRNA and/or degradation of aberrant  
 CC transcripts in a cell. Agents binding to the complex can be identified  
 CC and included in therapeutic compositions useful as above, and/or used to  
 CC modulate peptidyl transferase activity during translation in cells. They  
 CC are also useful to modulate the efficiency of translation termination of  
 CC mRNA at a nonsense codon and/or promote degradation of aberrant  
 CC transcripts in cells. The method can be used to identify agents/  
 CC compositions modulating binding to MT1, useful to identify genes.  
 CC Sequences MAY7813-817 represent protein fragments from yeast superfamily  
 CC group I helicases  
 CC  
 SO Sequence 415 AA;

Query Match 91.2%; Score 31; DB 3; Length 415;  
 Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLFERV 9  
 DB 193 LETSLFERV 201

RESULT 2  
 ABW01201  
 ID ABW01201 standard; protein; 415 AA.

AC ABW01201;  
 DT 15-JAN-2004 (first entry)

DE Saccharomyces cerevisiae modulator of translation termination protein.  
 KW Modulator of translation termination; MT1; helicase B; antiviral;  
 KM therapy; HCSB; nonsense mutation; yeast.

OS Saccharomyces cerevisiae.

XX US6630294-B1.

XX 07-OCT-2003.

XX 22-JUL-1999; 99US-00359268.

XX 22-JUL-1998; 98US-0093685P.

XX (UTNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

XX Peltz S, Czaplinski K, Dimman JD;

XX WPI; 2003-810549/76.

PT Identifying an agent that increases nonsense suppression, for antiviral  
 PT therapy, by contacting modulator of translation termination (Mtl1) in  
 PT Saccharomyces cerevisiae with a test agent, and detecting specific  
 PT binding to Mtl1.

XX Disclosure; Col 49-52; Opp; English.

XX The invention relates to a method of identifying an agent that increases  
 CC nonsense suppression, by contacting modulator of translation termination  
 CC (MT1) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.  
 CC The method is useful for identifying compositions or agents which  
 CC increase nonsense suppression. The invention may also be used for  
 CC antiviral therapy and for suppression of pathological nonsense mutations.

CC The present sequence is Saccharomyces cerevisiae MT1 protein  
 XX  
 SO Sequence 415 AA;

Query Match 91.2%; Score 31; DB 7; Length 415;  
 Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLFERV 9  
 DB 193 LETSLFERV 201

RESULT 3  
 ADP44127  
 ID ADP44127 standard; protein; 415 AA.

AC ADP44127;  
 DT 18-NOV-2004 (first entry)

DE Yeast helicase Mtl1.

XX gene therapy; translation termination; RNA helicase; MT1;  
 KW frameshift frequency; aberrant transcript degradation;  
 KW peptidyl transferase modulation; beta-thalassemia; beta-globin;  
 KW Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
 KW Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;  
 KW Ovarian Cancer; Wilms Tumor; Hirschsprung disease; Cystic fibrosis;  
 KW Kidney Stone; Familial hypercholesterolemia; Retinitis Pigmentosa;  
 KW Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; yeast; enzyme.

OS Saccharomyces cerevisiae.

XX US2004115787-A1.

XX 17-JUN-2004.

XX 28-AUG-2003; 2003US-00652334.

XX 22-JUL-1998; 98US-0093685P.

XX 22-JUL-1999; 99US-00359268.

XX (PELTZ) PELTZ S.  
 XX (CZAP/) CZAPLINSKI K.  
 XX (DINM/) DINMAN J D.

XX Peltz S, Czaplinski K, Dimman JD;

XX WPI; 2004-449400/42.

PT Identifying a test composition or agent that modulates the efficiency of  
 PT translation termination comprises contacting the MT1 with the test  
 PT composition or agent, and determining if the test composition or agent  
 PT inhibits the MT1.

XX Disclosure; SEQ ID NO 25; 41pp; English.

XX The invention relates to a method of identifying a test composition that  
 CC modulates the efficiency of translation termination comprising contacting  
 CC the RNA helicase MT1 with a composition or agent under conditions  
 CC permitting binding between the MT1 and the composition, detecting  
 CC specific binding of the test composition or agent to the MT1, and  
 CC determining if the test composition or agent inhibits the MT1. The  
 CC composition and methods are useful for modulating the fidelity of  
 CC translation termination or for identifying agents that affect the  
 CC functional activity of mRNAs by altering frameshift frequency, permit  
 CC monitoring of a termination event, promote degradation of aberrant  
 CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl  
 CC transferase activity during initiation, elongation, termination and mRNA  
 CC degradation of translation. The agents, which may be antagonists or  
 CC agonists, are useful in screening, diagnostic and therapeutic purposes,  
 CC for diseases or conditions resulting from or cause premature translation.

CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular  
 CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,  
 CC Osteogenesis Imperfecta, Breast Cancer, Ovarian Cancer, Wilms Tumour,  
 CC Hirschsprung disease, Cystic Fibrosis, Kidney Stones, Familial  
 CC hypercholesterolemia, Retinitis Pigmentosa, or Neurofibromatosis,  
 CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents  
 CC the amino acid sequence of the yeast helicase Mtt1.  
 XX

SO Sequence 415 AA;

Query Match 91.2%; Score 31; DB 8; Length 415;  
 Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LXXSLFPRV 9  
 DB 193 LETSLFPRV 201

RESULT 4  
 ABR01181  
 ID ABR01181 standard; peptide: 9 AA.  
 AC ABR01181;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Saccharomyces cerevisiae motif V peptide.  
 XX  
 KW Modulator of translation termination; MTT1; helicase B; antiviral;  
 KW therapy; HCSB; nonsense mutation; yeast.  
 XX  
 OS Saccharomyces cerevisiae.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 2 /label= Unknown  
 FT /note= "Xaa may be any amino acid"  
 FT Misc-difference 3 /label= Unknown  
 FT /note= "Xaa may be any amino acid"  
 XX  
 PN US6630294-B1.  
 XX  
 PD 07-OCT-2003.  
 XX  
 PF 22-JUL-1999; 99US-00359268.  
 XX  
 PR 22-JUL-1998; 98US-0093685P.  
 XX  
 PA (UTNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.  
 PI Peltz S, Czaplinski K, Dinman JD;  
 XX  
 DR WPI; 2003-810549/76.  
 XX  
 XX Identifying an agent that increases nonsense suppression, for antiviral  
 PT therapy, by contacting modulator of translation termination (Mtt1) in  
 PT Saccharomyces cerevisiae with a test agent, and detecting specific  
 PT binding to Mtt1.  
 XX  
 PS Disclosure; Col 43; Opp; English.  
 XX  
 CC The invention relates to a method of identifying an agent that increases  
 CC nonsense suppression, by contacting modulator of translation termination  
 CC (MTT1) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.  
 CC The method is useful for identifying compositions or agents which  
 CC increase nonsense suppression. The invention may also be used for  
 CC antiviral therapy and for suppression of pathological nonsense mutations.  
 CC The present sequence is Saccharomyces cerevisiae motif V peptide  
 XX  
 SO Sequence 9 AA;

Query Match 88.2%; Score 30; DB 7; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LXXSLFPRV 9  
 DB 1 LXXSLFPRV 9

RESULT 5  
 ADP44107  
 ID ADP44107 standard; peptide: 9 AA.  
 AC ADP44107;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Yeast translation termination modulation protein motif V.  
 XX  
 KW gene therapy; translation termination; RNA helicase; MTT1;  
 KW frameshift frequency; aberrant transcript degradation;  
 KW peptidyl transferase modulation; beta-thalassemia; beta-globin;  
 KW Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
 KW Von Willebrand Disease; Osteogenesis Imperfecta; Breast Cancer;  
 KW Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic Fibrosis;  
 KW Kidney Stone; Familial hypercholesterolemia; Retinitis Pigmentosa;  
 KW Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; yeast.  
 XX  
 OS Saccharomyces cerevisiae.  
 XX  
 FH Synthetic.  
 FT Key Location/Qualifiers  
 FT Misc-difference 2.3 /note= "Any amino acid"  
 FT

US2004115787-A1.  
 XX  
 PD 17-JUN-2004.  
 XX  
 PP 28-AUG-2003; 2003US-00652334.  
 XX  
 PR 22-JUL-1998; 98US-0093685P.  
 PR 22-JUL-1999; 99US-00359268.  
 XX  
 PA (PELTZ/) PELTZ S.  
 PA (CZAP/) CZAPINSKI K.  
 PA (DINM/) DINMAN J D.  
 PI Peltz S, Czaplinski K, Dinman JD;  
 XX  
 DR WPI; 2004-449400/42.  
 XX  
 XX Identifying a test composition or agent that modulates the efficiency of  
 PT translation termination comprises contacting the MTT1 with the test  
 PT composition or agent, and determining if the test composition or agent  
 PT inhibits the MTT1.  
 XX  
 PS Claim 37; SEQ ID NO 5; 41pp; English.  
 XX  
 CC The invention relates to a method of identifying a test composition that  
 CC modulates the efficiency of translation termination comprising contacting  
 CC the RNA helicase MTT1 with a composition or agent under conditions  
 CC permitting binding between the MTT1 and the composition, detecting  
 CC specific binding of the test composition or agent to the MTT1, and  
 CC determining if the test composition or agent inhibits the MTT1. The  
 CC composition and methods are useful for modulating the fidelity of  
 CC translation termination or for identifying agents that affect the  
 CC functional activity of mRNAs by altering frameshift frequency, permit  
 CC monitoring of a termination event, promote degradation of aberrant  
 CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl  
 CC transferase activity during initiation, elongation, termination and mRNA  
 CC degradation of translation. The agents, which may be antagonists or  
 CC agonists, are useful in screening, diagnostic and therapeutic purposes,

for diseases or conditions resulting from or cause premature translation, CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease, CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour, CC Hirschsprung disease, Cystic Fibrosis, Kidney Stones, Familial CC hypercholesterolaemia, Retinitis Pigmentosa, or Neurofibromatosis, CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents CC the amino acid sequence of the yeast translation termination modulation CC protein motif V.

XX  
SQ Sequence 9 AA;

Query Match 88.2%; Score 30; DB 8; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LXXSLFERV 9  
1 LXXSLFERV 9

DB 1 LXXSLFERV 9

RESULT 6  
AAV77808 standard; peptide; 10 AA.

XX  
AC AAV77808;  
XX  
DT 31-MAY-2000 (first entry)  
XX  
DE Motif V comprised in a gene modulating translation termination.  
XX  
KM Helicase B; HCSB; MTT1; modulator of translation termination; eRF1; eRF3;  
KM eukaryotic release factor; peptidyl transferase; beta-thalassemia;  
KM beta-globin; Duchene/Becker Muscular Dystrophy; anti-anemic.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1..10 /note= "residues indicated Xaa are unspecified"

XX  
PN W0200005586-A2.  
XX  
PD 03-FEB-2000.  
XX  
PF 22-JUL-1999; 99WO-US016802.  
XX  
PR 22-JUL-1998; 98US-00120435.  
XX  
PA (UYNE-) UNIV NEW JERSEY.  
XX  
PI Peltz S, Czaplinski K, Dimman JD;  
XX  
DR WPI; 2000-171458/15.  
XX  
PT New multiprotein complex which can modulate peptidyl transferase activity  
PT during translation, useful to treat diseases associated with peptidyl  
PT transferase activity e.g. Duchene/Becker Muscular Dystrophy.  
XX  
PS Claim 37; Page 79; 89pp; English.  
XX  
CC The invention provides a new multiprotein complex which can modulate  
CC peptidyl transferase activity during translation. The complex comprises  
CC the gene encoding Helicase B (HCSB; renamed MTT1, for Modulator of  
CC translation Termination) and the conserved proteins known to interact and  
CC carry out translation termination in eukaryotic cells, peptidyl  
CC eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to  
CC modulate peptidyl transferase activity during translation in a cell. It  
CC can be administered therapeutically combined with a carrier in  
CC pharmaceutical compositions to treat diseases associated with peptidyl  
CC transferase activity, especially diseases resulting from a nonsense or  
CC frameshift mutation e.g. beta-thalassemia, beta-globin, Duchene/Becker  
CC Muscular Dystrophy etc. It can be used to identify disease conditions

CC involving a defect in the complex, by transfecting cells with encoding  
CC nucleic acid and determining the proportion of defective complex before  
CC and after transfection. It is also useful to screen for drugs involved in  
CC peptidyl transferase activity during translation, inhibiting the  
CC interaction between MTT1 and eRF3 or involved in enhancing translation  
CC termination. Vectors comprising polynucleotides encoding the complex (or  
CC antisense sequences) can be constructed and introduced into cells to  
CC interfere with complex expression and so modulate the efficiency of  
CC translation termination of mRNA and/or degradation of aberrant  
CC transcripts in a cell. Agents binding to the complex can be identified  
CC and included in therapeutic compositions useful as above, and/or used to  
CC modulate peptidyl transferase activity during translation in cells. They  
CC are also useful to modulate the efficiency of translation termination of  
CC mRNA at a nonsense codon and/or promote degradation of aberrant  
CC transcripts in cells. The method can be used to identify agents/  
CC compositions modulating binding to MTT1, useful to identify genes.  
CC Sequences AAV77804-812 represent motifs I-IX comprised in the genes of  
CC interest, used for modulating translation termination

XX  
SQ Sequence 10 AA;

Query Match 88.2%; Score 30; DB 3; Length 10;  
Best Local Similarity 100.0%; Pred. No. 5.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LXXSLFERV 9  
1 LXXSLFERV 9

DB 1 LXXSLFERV 9

RESULT 7  
ABW01186 standard; peptide; 10 AA.

XX  
ID ABW01186  
XX  
AC ABW01186;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Saccharomyces cerevisiae consensus motif peptide #1.  
XX  
KM Modulator of translation termination; MTT1; helicase B; antiviral;  
KM therapy; HCSB; nonsense mutation; yeast.  
XX  
OS Saccharomyces cerevisiae.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 3 /label= Unknown  
FT /note= "Xaa may be any amino acid"

XX  
PN US6630294-B1.  
XX  
PD 07-OCT-2003.  
XX  
PF 22-JUL-1999; 99US-00359268.  
XX  
PR 22-JUL-1998; 98US-0093685P.  
XX  
PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.  
XX  
PI Peltz S, Czaplinski K, Dimman JD;  
XX  
DR WPI; 2003-810549/76.  
XX  
PT Identifying an agent that increases nonsense suppression; for antiviral  
PT therapy, by contacting modulator of translation termination (Mtt1) in  
PT Saccharomyces cerevisiae with a test agent, and detecting specific  
PT binding to Mtt1.  
XX  
PS Example; Col 45; Opp; English.  
XX  
CC The invention relates to a method of identifying an agent that increases  
CC nonsense suppression, by contacting modulator of translation termination

CC (MTT1) also referred to as helicase B (HCSB) in *Saccharomyces cerevisiae*.  
CC The method is useful for identifying compositions or agents which  
CC increase nonsense suppression. The invention may also be used for  
CC antiviral therapy and for suppression of pathological nonsense mutations.  
CC The present sequence is *Saccharomyces cerevisiae* consensus motif peptide  
XX  
SQ Sequence 10 AA;  
Query Match 88.2%; Score 30; DB 7; Length 10;  
Best Local Similarity 88.9%; Pred. No. 5.7;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 LXXSLPERV 9  
1 LXXSLPERV 9  
1 LXXSLPERV 9  
Db  
RESULT 8  
ADP44112  
ID ADP44112 standard; peptide; 10 AA.  
XX  
AC ADP44112;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Yeast translation termination modulation protein related peptide #1.  
XX  
KM gene therapy; translation termination; RNA helicase; MTT1;  
KM frameshift frequency; aberrant transcript degradation;  
KM peptidyl transferase modulation; beta-thalassemia; beta-globin;  
KM Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
KM Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;  
KM Ovarian Cancer; Wilms Tumour; Hirschprung disease; Cystic fibrosis;  
KM Kidney Stone; Familial hypercholesterolemia; Retinitis Pigmentosa;  
KM Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; Yeast.  
XX  
OS *Saccharomyces cerevisiae*.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 3 /note= "Any amino acid"  
XX  
PN US2004115787-A1.  
XX  
PD 17-JUN-2004.  
XX  
PP 28-AUG-2003; 2003US-00652334.  
XX  
PR 22-JUL-1998; 98US-0093685P.  
PR 22-JUL-1999; 99US-00359268.  
XX  
PA (PELTZ/) PELTZ S.  
PA (CZAP/) CZAPLINSKI K.  
PA (DINM/) DINMAN J D.  
XX  
PI Pelitz S, Czaplinski K, Dinman JD;  
XX  
DR WPI; 2004-449400/42.  
XX  
PT Identifying a test composition or agent that modulates the efficiency of  
PT translation termination comprises contacting the MTT1 with the test  
PT composition or agent, and determining if the test composition or agent  
PT inhibits the MTT1.  
XX  
PS Disclosure; SEQ ID NO 10; 41pp; English.  
XX  
CC The invention relates to a method of identifying a test composition that  
CC modulates the efficiency of translation termination comprising contacting  
CC the RNA helicase MTT1 with a composition or agent under conditions  
CC permitting binding between the MTT1 and the composition, detecting  
CC specific binding of the test composition or agent to the MTT1, and  
CC determining if the test composition or agent inhibits the MTT1. The  
CC composition and methods are useful for modulating the fidelity of

CC translation termination or for identifying agents that affect the  
CC functional activity of mRNAs by altering frameshift frequency, permit  
CC monitoring of a termination event, promote degradation of aberrant  
CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl  
CC transferase activity during initiation, elongation, termination and mRNA  
CC degradation of translation. The agents, which may be antagonists or  
CC agonists, are useful in screening, diagnostic and therapeutic purposes,  
CC for diseases or conditions resulting from or cause premature translation,  
CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular  
CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,  
CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,  
CC Hirschprung disease, Cystic fibrosis, Kidney Stones, Familial  
CC hypercholesterolemia, Retinitis Pigmentosa, or Neurofibromatosis,  
CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents  
CC the amino acid sequence of a Yeast translation termination modulation  
CC protein related peptide. Note: this sequence appears in the sequence  
CC listing but no reference is made to it in the main body of the  
CC specification.  
XX  
SQ Sequence 10 AA;  
Query Match 88.2%; Score 30; DB 8; Length 10;  
Best Local Similarity 88.9%; Pred. No. 5.7;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 LXXSLPERV 9  
1 LXXSLPERV 9  
1 LXXSLPERV 9  
Db  
RESULT 9  
ABM01195  
ID ABM01195 standard; peptide; 7 AA.  
XX  
AC ABM01195;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE *Saccharomyces cerevisiae* consensus peptide #8.  
XX  
KM Modulator of translation termination; MTT1; helicase B; antiviral;  
KM therapy; HCSB; nonsense mutation; Yeast.  
XX  
OS *Saccharomyces cerevisiae*.  
XX  
PN US6630294-B1.  
XX  
PD 07-OCT-2003.  
XX  
PP 22-JUL-1999; 99US-00359268.  
XX  
PR 22-JUL-1998; 98US-0093685P.  
XX  
PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.  
XX  
PI Pelitz S, Czaplinski K, Dinman JD;  
XX  
DR WPI; 2003-810549/76.  
XX  
PT Identifying an agent that increases nonsense suppression, for antiviral  
PT therapy, by contacting modulator of translation termination (Mtt1) in  
PT *Saccharomyces cerevisiae* with a test agent, and detecting specific  
PT binding to Mtt1.  
XX  
PS Disclosure; Col 47; 0pp; English.  
XX  
CC The invention relates to a method of identifying an agent that increases  
CC nonsense suppression, by contacting modulator of translation termination  
CC (MTT1) also referred to as helicase B (HCSB) in *Saccharomyces cerevisiae*.  
CC The method is useful for identifying compositions or agents which  
CC increase nonsense suppression. The invention may also be used for  
CC antiviral therapy and for suppression of pathological nonsense mutations.  
CC The present sequence is *Saccharomyces cerevisiae* consensus peptide

XX Sequence 7 AA;  
SQ

Query Match 82.4%; Score 28; DB 7; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLFPRV 9  
|||  
1 SLFPRV 6

Db

RESULT 10  
ADP44121  
ID ADP44121 standard; peptide; 7 AA.

XX ADP44121;  
AC  
DT 18-NOV-2004 (first entry)  
XX

DE Yeast translation termination modulation protein related peptide #10.

XX gene therapy; translation termination; RNA helicase; MTT1;  
KW frameshift frequency; aberrant transcript degradation;  
KW peptidyl transferase modulation; beta-thalassemia; beta-globin;  
KW Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
KW Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;  
KW Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic fibrosis;  
KW Kidney Stone; Familial hypercholesterolemia; Retinitis Pigmentosa;  
KW Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; Yeast.

XX Saccharomyces cerevisiae.  
XX

PN US2004115787-A1.  
PD 17-JUN-2004.  
PF 28-AUG-2003; 2003US-00652334.  
PR 22-JUL-1998; 98US-0093688P.  
PR 22-JUL-1999; 99US-00359288.  
XX

PA (PELTZ/) PELTZ S.  
PA (CZAP/) CZAPLINSKI K.  
PA (DINM/) DINMAN J D.  
XX

PI Pelcz S, Czaplinski K, Dinman JD;  
XX

DR WPI; 2004-449400/42.  
XX

PT Identifying a test composition or agent that modulates the efficiency of  
PT translation termination comprises contacting the MTT1 with the test  
PT composition or agent, and determining if the test composition or agent  
PT inhibits the MTT1.  
XX

PS Disclosure; SEQ ID NO 19; 41pp; English.  
XX

XX The invention relates to a method of identifying a test composition that  
XX modulates the efficiency of translation termination comprising contacting  
XX the RNA helicase MTT1 with a composition or agent under conditions  
XX permitting binding between the MTT1 and the composition, detecting  
XX specific binding of the test composition or agent to the MTT1, and  
XX determining if the test composition or agent inhibits the MTT1. The  
XX composition and methods are useful for modulating the fidelity of  
XX translation termination or for identifying agents that: affect the  
XX functional activity of mRNAs by altering frameshift frequency, permit  
XX monitoring of a termination event, promote degradation of aberrant  
XX transcripts, and provide modulators (inhibitors/stimulators) of peptidyl  
XX transferase activity during initiation, elongation, termination and mRNA  
XX degradation of translation. The agents, which may be antagonists or  
XX agonists, are useful in screening, diagnostic and therapeutic purposes,  
XX for diseases or conditions resulting from or cause premature translation,  
XX such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular

CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,  
CC Osteogenesis Imperfecta, Breast Cancer, Ovarian Cancer, Wilms Tumour,  
CC Hirschsprung disease, Cystic fibrosis, Kidney Stones, Familial  
CC hypercholesterolemia, Retinitis Pigmentosa, or Neurofibromatosis,  
CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents  
CC the amino acid sequence of a Yeast translation termination modulation  
CC protein related peptide. Note: this sequence appears in the sequence  
CC listing but no reference is made to it in the main body of the  
CC specification.  
XX

XX Sequence 7 AA;  
SQ

Query Match 82.4%; Score 28; DB 8; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLFPRV 9  
|||  
1 SLFPRV 6

Db

RESULT 11  
ABO58160  
ID ABO58160 standard; protein; 71 AA.

AC ABO58160;  
XX

DT 29-JUL-2004 (first entry)  
XX

DE Human genome derived single exon protein #4394.  
XX

XX Human; gene expression; single exon probe; microarray;  
KW alternative splicing event; genomic alteration.  
XX

OS Homo sapiens.  
XX

PN US2003194704-A1.  
PD 16-OCT-2003.  
PF 03-APR-2002; 2002US-00029386.  
PR 03-APR-2002; 2002US-00029386.  
XX

PA (PENN/) PENN S G.  
PA (RANK/) RANK D R.  
PA (HANZ/) HANZEL D K.  
XX

PI Penn SG, Rank DR, Hanzel DK;  
XX

DR WPI; 2004-119264/12.  
XX

PT New human genome-derived single exon nucleic acid probes useful for human  
PT gene expression analysis, for identifying or characterizing alternative  
PT splicing events, for assessing genomic alterations or as tools for  
PT surveying tissues.  
XX

PS Claim 45; SEQ ID NO 31794; 80pp; English.  
XX

XX The invention relates to a nucleic acid probe for measuring human gene  
XX expression, comprising any of the 27,400 fully defined nucleotide  
XX sequences in the specification, or their complements or fragments, and  
XX encoding at least 8 amino acids of any of the 688 amino acid sequences  
XX fully defined in the specification. The probe is a single exon probe that  
XX hybridises under high stringency conditions to a nucleic acid molecule  
XX expressed in human cells or tissues. Also included are a spatially-  
XX addressable set of single exon nucleic acid probes for measuring human  
XX gene expression (comprising a plurality of single exon nucleic acid  
XX probes cited above, where each of the plurality of probes is separately  
XX and addressably isolatable or amplifiable from the plurality), a single  
XX exon microarray for measuring human gene expression, a method of  
XX measuring human gene expression, a vector comprising the single exon  
XX probe cited above, an ORF-encoded peptide comprising at least 8

CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics  
CC and treatment of pathological conditions resulting from a  
CC prophylaxis and treatment of pathological conditions resulting from a

CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The



CC sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html

XX Sequence 260 AA;

Query Match 82.4%; Score 28; DB 7; Length 260;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLFPRV 9  
 DB 231 SLFPRV 236

RESULT 14  
 ABO63671  
 ID ABO63671 standard; protein, 298 AA.

XX ABO63671;

XX 29-JUN-2004 (first entry)

XX Klebsiella pneumoniae polypeptide seqid 10188.

XX Recombinant expression vector; transcription regulatory element;

KW Klebsiella pneumoniae protein; antibacterial; Vaccine.

XX Klebsiella pneumoniae.

XX US6610836-B1.

XX 26-AUG-2003.

XX 27-JAN-2000; 2000US-00489039.

XX 29-JAN-1999; 99US-0117747P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Breton GL, Osborne M;

XX WPI; 2003-895346/82.

DR N-PSDB; ACH97222.

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for  
 PT preparing a vaccine composition against Klebsiella pneumoniae.

PS Disclosure; SEQ ID NO 10188; 932pp; English.

XX The invention describes a new isolated nucleic acid encoding a Klebsiella  
 CC pneumoniae polypeptide. Also described are: a recombinant expression  
 CC vector comprising the nucleic acid, operably linked to a transcription  
 CC regulatory element; and a cell comprising the recombinant expression  
 CC vector. The nucleic acid is useful for preparing a vaccine composition  
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a  
 CC Klebsiella pneumoniae polypeptide of the invention

XX Sequence 298 AA;

Query Match 82.4%; Score 28; DB 7; Length 298;  
 Best Local Similarity 75.0%; Pred. No. 4.3e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LXXSLFER 8  
 DB 48 LSASLFER 55

RESULT 15  
 ABUS2076  
 ID ABUS2076 standard; protein, 315 AA.

AC ABUS2076;

XX 07-MAY-2003 (first entry)

XX Helicobacter pylori selected interacting domain (SID) protein #1420.

XX Protein-protein interaction; ulcer; selected interacting domain; SID.

KW Helicobacter pylori.

XX WO200266501-A2.

XX 29-AUG-2002.

XX 28-DEC-2001; 2001WO-EP015428.

XX 02-JAN-2001; 2001US-0259302P.

XX (HYBR-) HYBRIGENICS.

PA (INSP ) INST PASTEUR.

PI Legrain P, Rain J, Collard F, De Reuse H, Labigne A;

XX WPI; 2002-674910/72.

DR N-PSDB; ABX66821.

XX New complexes of protein-protein interactions in Helicobacter pylori,  
 PT useful for identifying modulating compounds for treating or preventing  
 PT ulcers in mammals.

XX Claim 6, Page 431; 642pp; English.

XX The invention describes a complex of protein-protein interactions in  
 CC Helicobacter pylori selected from 421 complexes given in the  
 CC specification. The complex of protein-protein interactions are useful for  
 CC screening for agents which mediate the interaction of proteins.  
 CC Modulating compounds which binds to a targeted bacterial protein may be  
 CC used for treating or preventing ulcers in a human or animal. This is the  
 CC amino acid sequence of a selected interacting domain (SID), identified  
 CC via protein-protein interactions. Note: Where the patent number printed  
 CC at the top of the pages in the specification has obscured areas of  
 CC protein sequence, the indexer has replaced the residue with an X to  
 CC represent an illegible residue

XX Sequence 315 AA;

Query Match 82.4%; Score 28; DB 5; Length 315;  
 Best Local Similarity 66.7%; Pred. No. 4.5e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LXXSLFERV 9  
 DB 181 LSLSLPRV 189

Search completed: April 18, 2005, 08:03:48  
 Job time : 46.914 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using BW model

Run on: April 18, 2005, 07:37:21 ; Search time 11.2336 Seconds  
(without alignments)  
59.807 Million cell updates/sec

Title: US-10-652-334-5  
34  
Perfect score: 1 LXXSLPERV 9  
Sequence: 1 LXXSLPERV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/6CTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	91.2	415	US-09-359-268A-25	Sequence 25, Appl
2	31	91.2	917	US-09-248-796A-19347	Sequence 19347, A
3	30	88.2	9	US-09-359-268A-5	Sequence 5, Appl
4	30	88.2	10	US-09-359-268A-10	Sequence 10, Appl
5	28	82.4	7	US-09-359-268A-19	Sequence 19, Appl
6	28	82.4	195	US-09-902-540-11930	Sequence 11930, A
7	28	82.4	230	US-09-252-991A-27621	Sequence 27621, A
8	28	82.4	260	US-09-252-991A-29356	Sequence 29356, A
9	28	82.4	298	US-09-489-039A-10188	Sequence 10188, A
10	28	82.4	1043	US-08-724-354D-4	Sequence 4, Appl
11	28	82.4	1043	US-09-270-984A-4	Sequence 4, Appl
12	28	82.4	1118	US-08-724-354D-2	Sequence 2, Appl
13	28	82.4	1118	US-09-270-984A-2	Sequence 2, Appl
14	28	82.4	1140	US-09-949-016-10116	Sequence 10116, A
15	28	82.4	1240	US-08-930-996A-4	Sequence 4, Appl
16	27	79.4	106	US-09-248-796A-18134	Sequence 18134, A
17	27	79.4	127	US-09-902-540-14345	Sequence 14345, A
18	27	79.4	147	US-09-543-681A-5609	Sequence 5609, Ap
19	27	79.4	380	US-09-252-991A-17188	Sequence 17188, A
20	27	79.4	380	US-09-359-268A-29	Sequence 29, Appl
21	27	79.4	498	US-09-248-796A-16174	Sequence 16174, A
22	27	79.4	718	US-09-252-991A-33109	Sequence 33109, A
23	27	79.4	971	US-08-724-354D-22	Sequence 22, Appl
24	27	79.4	971	US-09-270-984A-22	Sequence 22, Appl
25	27	79.4	971	US-09-177-431-8	Sequence 8, Appl
26	27	79.4	1411	US-09-252-991A-28408	Sequence 28408, A
27	27	79.4	200	US-09-134-000C-5989	Sequence 5989, Ap

28	26	76.5	203	4	US-09-489-039A-9971	Sequence 9971, Ap
29	26	76.5	284	4	US-09-252-991A-26168	Sequence 26168, A
30	26	76.5	323	4	US-09-522-714-8	Sequence 8, Appl
31	26	76.5	324	4	US-09-905-290A-4	Sequence 4, Appl
32	26	76.5	546	4	US-09-543-681A-7065	Sequence 7065, Ap
33	26	76.5	556	4	US-09-902-540-13058	Sequence 13058, A
34	26	76.5	970	4	US-09-134-000C-4256	Sequence 4256, Ap
35	26	76.5	999	2	US-08-770-301A-1	Sequence 1, Appl
36	26	76.5	1506	4	US-09-175-581-1	Sequence 1, Appl
37	26	76.5	1506	4	US-09-902-540-10944	Sequence 10944, A
38	25	73.5	61	4	US-09-248-796A-25656	Sequence 25656, A
39	25	73.5	102	4	US-09-270-767-58044	Sequence 58044, A
40	25	73.5	126	4	US-09-270-767-57532	Sequence 37532, A
41	25	73.5	126	4	US-09-270-767-52749	Sequence 52749, A
42	25	73.5	139	4	US-09-252-991A-22668	Sequence 22668, A
43	25	73.5	152	4	US-09-270-767-35592	Sequence 35592, A
44	25	73.5	152	4	US-09-270-767-42726	Sequence 42726, A
45	25	73.5	152	4	US-09-270-767-50809	Sequence 50809, A

## ALIGNMENTS

```
RESULT 1
US-09-359-268A-25
; Sequence 25, Application US/09359268A
; Patent No. 6630294
; GENERAL INFORMATION:
; APPLICANT: Peltz, Stuart
; APPLICANT: Czaplinski, Kevin
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
; FILE REFERENCE: 601-1-85N
; CURRENT APPLICATION NUMBER: US/09/359,268A
; CURRENT FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,685
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 25
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-359-268A-25

Query Match          91.2%  Score 31;  DB 4;  Length 415;
Best Local Similarity 77.8%  Pred. No. 23;
Matches 7;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;

Cy      1 LXXSLPERV 9
Db      193 LETSLPERV 201

RESULT 2
US-09-248-796A-19347
; Sequence 19347, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19347
```

LENGTH: 917  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-19347

Query Match 91.2%; Score 31; DB 4; Length 917;  
Best Local Similarity 77.8%; Pred. No. 52;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLFERV 9  
Db 662 LSLSLFERV 670

RESULT 3  
US-09-359-268A-5  
Sequence 5, Application US/09359268A  
Patent No. 6630294  
GENERAL INFORMATION:  
APPLICANT: Peltz, Stuart  
APPLICANT: Czaplinski, Kevin  
APPLICANT: Diman, Jonathan D.  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/09/359,268A  
PRIOR FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 5  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
FEATURE:  
OTHER INFORMATION: Xaa= any amino acid  
US-09-359-268A-5

Query Match 88.2%; Score 30; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LXXSLFERV 9  
Db 1 LXXSLFERV 9

RESULT 4  
US-09-359-268A-10  
Sequence 10, Application US/09359268A  
Patent No. 6630294  
GENERAL INFORMATION:  
APPLICANT: Peltz, Stuart  
APPLICANT: Czaplinski, Kevin  
APPLICANT: Diman, Jonathan D.  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/09/359,268A  
PRIOR FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 10  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
FEATURE:  
OTHER INFORMATION: Xaa = any amino acid

US-09-359-268A-10

Query Match 88.2%; Score 30; DB 4; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.8;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LXXSLFERV 9  
Db 1 LXXSLFERV 9

RESULT 5  
US-09-359-268A-19  
Sequence 19, Application US/09359268A  
Patent No. 6630294  
GENERAL INFORMATION:  
APPLICANT: Peltz, Stuart  
APPLICANT: Czaplinski, Kevin  
APPLICANT: Diman, Jonathan D.  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/09/359,268A  
PRIOR FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 19  
LENGTH: 7  
TYPE: PRT  
ORGANISM: saccharomyces cerevisiae  
US-09-359-268A-19

Query Match 82.4%; Score 28; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLFERV 9  
Db 1 SLFERV 6

RESULT 6  
US-09-902-540-11930  
Sequence 11930, Application US/09902540  
Patent No. 6833447  
GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Wiegand, Roger C.  
TITLE OF INVENTION: MYXOCOCCUS XANTHUS Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(115849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 11930  
LENGTH: 195  
TYPE: PRT  
ORGANISM: Myxococcus xanthus  
US-09-902-540-11930

Query Match 82.4%; Score 28; DB 4; Length 195;  
Best Local Similarity 66.7%; Pred. No. 52;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLFERV 9  
Db 73 LAOTLFERV 81

RESULT 7  
US-09-252-991A-27621  
; Sequence 27621, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 27621  
; LENGTH: 230  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27621

Query Match 82.4%; Score 28; DB 4; Length 230;  
Best Local Similarity 66.7%; Pred. No. 62;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LXXSLPERV 9  
| | | | |  
Db 208 LLSALPERV 216

RESULT 8  
US-09-252-991A-29356  
; Sequence 29356, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 29356  
; LENGTH: 260  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29356

Query Match 82.4%; Score 28; DB 4; Length 260;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLPERV 9  
| | | | |  
Db 231 SLPERV 236

RESULT 9  
US-09-489-039A-10188  
; Sequence 10188, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 10188  
; LENGTH: 298  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-10188

Query Match 82.4%; Score 28; DB 4; Length 298;  
Best Local Similarity 75.0%; Pred. No. 81;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LXXSLPER 8  
| | | | |  
Db 48 LLSALPER 55

RESULT 10  
US-08-724-354D-4  
; Sequence 4, Application US/08724354D  
; Patent No. 5994119  
; GENERAL INFORMATION:  
; APPLICANT: Dietz, Harry C.  
; TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
; TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Fish & Richardson, P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: IBM Compatible  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,354D  
; FILING DATE: 01-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/016,482  
; FILING DATE: 29-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hallie, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/090001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-678-5070  
; TELEFAX: 619-678-5099  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1043 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-724-354D-4

Query Match 82.4%; Score 28; DB 2; Length 1043;  
Best Local Similarity 66.7%; Pred. No. 3e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LXXSLPERL 9  
| | | | |  
Db 603 LLSLPERL 611

RESULT 11  
US-09-270-984A-4

```
; Sequence 4, Application US/09270984A
; Patent No. 6048965
; GENERAL INFORMATION:
; APPLICANT: Dietz, Harry C.
; TITLE OF INVENTION: MAMMALIAN REGULATOR OF
; TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECA
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/270,984A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/724,354
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/090001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5099
; TELEFAX: 619-678-5099
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1043 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-270-984A-4

Query Match      82.4%; Score 28; DB 3; Length 1043;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/016,482
; FILING DATE: 29-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/090001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5099
; TELEFAX: 619-678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1118 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-724-354D-2

Query Match      82.4%; Score 28; DB 2; Length 1118;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
RESULT 13
US-09-270-984A-2
; Sequence 2, Application US/09270984A
; Patent No. 6048965
; GENERAL INFORMATION:
; APPLICANT: Dietz, Harry C.
; TITLE OF INVENTION: MAMMALIAN REGULATOR OF
; TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECA
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/270,984A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/724,354
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/090001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5099
; TELEFAX: 619-678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1118 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-270-984A-2

Query Match      82.4%; Score 28; DB 3; Length 1118;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

QY 1 LXXSLPERV 9  
Db 680 LSGSLPERL 688

RESULT 14  
US-09-949-016-10116  
Sequence 10116, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: C1001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10116  
LENGTH: 1140  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-10116

Query March 82.4%; Score 28; DB 4; Length 1140;  
Best Local Similarity 66.7%; Pred. No. 3.3e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LXXSLPERV 9  
Db 702 LSGSLPERL 710

RESULT 15  
US-08-930-996A-4  
Sequence 4, Application US/08930996A  
Patent No. 6100449  
GENERAL INFORMATION:  
APPLICANT: FLUHR, Robert  
APPLICANT: ESHED, Yuval  
APPLICANT: ORI, Naom1  
APPLICANT: PARAN, Ilan  
APPLICANT: ZAMIR, Daniel  
TITLE OF INVENTION: A GENE FAMILY FROM THE 12 FUSARIUM RESISTANCE  
LOCUS OF TOMATO AND USE THEREOF FOR TRANSFORMATION AND  
SELECTIVE BREEDING OF TOMATO AND RELATED PLANTS  
TITLE OF INVENTION: LOCUS OF TOMATO AND USE THEREOF FOR TRANSFORMATION AND  
SELECTIVE BREEDING OF TOMATO AND RELATED PLANTS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/930,996A  
FILING DATE: 09-DEC-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/05272  
FILING DATE: 15-APR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IL 113,373

FILING DATE: 13-APR-1995  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1240 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-930-996A-4

Query March 82.4%; Score 28; DB 3; Length 1240;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLPERV 9  
Db 472 SLPERV 477

Search completed: April 18, 2005, 08:18:34  
Job time: 11.2336 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 08:06:16 ; Search time 31.927 Seconds  
(without alignments)  
93.693 Million cell updates/sec

Title: US-10-652-334-5  
Perfect score: 34  
Sequence: 1 LXXSLFERV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	31	91.2	415	US-10-652-334-25
2	30	88.2	9	US-10-652-334-5
3	30	88.2	10	US-10-652-334-10
4	28	82.4	7	US-10-652-334-19
5	28	82.4	71	US-10-029-386-31794
6	28	82.4	199	US-10-767-701-38822
7	28	82.4	638	US-10-437-963-181690
8	28	82.4	734	US-10-424-599-205643
9	28	82.4	821	US-10-369-493-18631
10	28	82.4	925	US-10-369-493-2087
11	28	82.4	935	US-10-408-765A-1706
12	28	82.4	1118	US-10-474-553-6
13	27	79.4	210	US-10-282-122A-67256

14	27	79.4	286	16	US-10-437-963-142655	Sequence 142655, A
15	27	79.4	298	15	US-10-282-122A-66301	Sequence 66301, A
16	27	79.4	304	16	US-10-437-963-173847	Sequence 173847, A
17	27	79.4	333	9	US-09-815-242-11916	Sequence 11916, A
18	27	79.4	333	15	US-10-282-122A-66479	Sequence 66479, A
19	27	79.4	333	15	US-10-389-647-571	Sequence 571, App
20	27	79.4	380	16	US-10-652-334-29	Sequence 29, Appl
21	27	79.4	413	15	US-10-282-122A-48431	Sequence 48431, A
22	27	79.4	604	15	US-10-282-122A-58045	Sequence 58045, A
23	27	79.4	644	15	US-10-335-977-8439	Sequence 8439, App
24	27	79.4	651	15	US-10-335-977-8440	Sequence 8440, App
25	27	79.4	839	16	US-10-032-585-7124	Sequence 7124, App
26	27	79.4	890	16	US-10-437-963-118530	Sequence 118530, A
27	27	79.4	971	15	US-10-369-493-11715	Sequence 11715, App
28	27	79.4	971	16	US-10-451-467A-698	Sequence 698, App
29	27	79.4	997	15	US-10-369-493-12790	Sequence 12790, A
30	27	79.4	2462	16	US-10-437-963-114113	Sequence 114113, A
31	27	79.4	2593	16	US-10-437-963-114115	Sequence 114115, A
32	26	76.5	77	15	US-10-424-599-180332	Sequence 180332, A
33	26	76.5	124	15	US-10-282-122A-61519	Sequence 61519, A
34	26	76.5	158	15	US-10-369-493-9883	Sequence 9883, App
35	26	76.5	175	16	US-10-767-701-46637	Sequence 46637, A
36	26	76.5	269	16	US-10-767-701-43960	Sequence 43960, A
37	26	76.5	273	16	US-10-767-701-44666	Sequence 44666, A
38	26	76.5	281	16	US-10-437-963-141966	Sequence 141966, A
39	26	76.5	296	16	US-10-437-963-153126	Sequence 153126, A
40	26	76.5	308	15	US-10-282-122A-49263	Sequence 49263, A
41	26	76.5	312	16	US-10-767-701-46662	Sequence 46662, A
42	26	76.5	314	16	US-10-437-963-162020	Sequence 162020, A
43	26	76.5	324	14	US-10-304-928-8	Sequence 8, Appl1
44	26	76.5	324	10	US-09-905-290A-4	Sequence 4, Appl1
45	26	76.5	348	16	US-10-437-963-132857	Sequence 132857, A

## ALIGNMENTS

RESULT 1  
US-10-652-334-25  
Sequence 25, Application US/10652334  
Publication No. US20040115787A1  
GENERAL INFORMATION:  
APPLICANT: Felitz, Stuart  
APPLICANT: Czaplinski, Kevin  
APPLICANT: Dimman, Jonathan D.  
TITLE OF INVENTION: A SUBRANITY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/10/652,334  
CURRENT FILING DATE: 2003-08-28  
PRIOR APPLICATION NUMBER: US/09/359,268A  
PRIOR FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
PRIOR FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patent Ver. 2.0  
SEQ ID NO 25  
LENGTH: 415  
TYPE: PRT  
ORGANISM: saccharomyces cerevisiae  
US-10-652-334-25

Query Match 91.2%; Score 31; DB 16; Length 415;  
Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 LXXSLFERV 9  
Db 193 LXXSLFERV 201

RESULT 2

```
US-10-652-334-5
; Sequence 5, Application US/10652334
; Publication No. US20040115787A1
; GENERAL INFORMATION:
; APPLICANT: Peltz, Stuart
; APPLICANT: Czaplinski, Kevin
; APPLICANT: Dimman, Jonathan D.
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 601-1-85N
; CURRENT APPLICATION NUMBER: US/10/652,334
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/359,268A
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,685
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; OTHER INFORMATION: Xaa = any amino acid
US-10-652-334-5

Query Match      88.2%; Score 30; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LXXSLPERV 9
DB 1 LXXSLPERV 9

RESULT 3
US-10-652-334-10
; Sequence 10, Application US/10652334
; Publication No. US20040115787A1
; GENERAL INFORMATION:
; APPLICANT: Peltz, Stuart
; APPLICANT: Czaplinski, Kevin
; APPLICANT: Dimman, Jonathan D.
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 601-1-85N
; CURRENT APPLICATION NUMBER: US/10/652,334
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/359,268A
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,685
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; OTHER INFORMATION: Xaa = any amino acid
US-10-652-334-10

Query Match      88.2%; Score 30; DB 16; Length 10;
Best Local Similarity 88.9%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
US-10-652-334-19
; Sequence 19, Application US/10652334
; Publication No. US20040115787A1
; GENERAL INFORMATION:
; APPLICANT: Peltz, Stuart
; APPLICANT: Czaplinski, Kevin
; APPLICANT: Dimman, Jonathan D.
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 601-1-85N
; CURRENT APPLICATION NUMBER: US/10/652,334
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/359,268A
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,685
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-652-334-19

Query Match      82.4%; Score 28; DB 16; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLPERV 9
DB 1 SLPERV 6

RESULT 5
US-10-029-386-31794
; Sequence 31794, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 31794
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003972.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3
; OTHER INFORMATION: SWISSPROT HIT: Q09820, EVALU 2.00e-24
US-10-029-386-31794

Query Match      82.4%; Score 28; DB 14; Length 71;
Best Local Similarity 66.7%; Pred. No. 91;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```



RESULT 6  
US-10-767-701-38822  
; Sequence 38822, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 38822  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28WAY03-C7426\_1.pep  
US-10-767-701-38822

Query Match 82.4%; Score 28; DB 16; Length 199;  
Best Local Similarity 66.7%; Pred. No. 2.5e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LKXSLPERV 9  
Db 69 LAOSLPERL 77

RESULT 7  
US-10-437-963-181690  
; Sequence 181690, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbasuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 181690  
; LENGTH: 638  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(638)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_7893C.1.pep  
US-10-437-963-181690

Query Match 82.4%; Score 28; DB 16; Length 638;  
Best Local Similarity 66.7%; Pred. No. 8e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LKXSLPERV 9  
Db 76 LAOSLPERL 84

RESULT 8  
US-10-424-599-205643

; Sequence 205643, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 205643  
; LENGTH: 734  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(734)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_27723C.1.pep  
US-10-424-599-205643

Query Match 82.4%; Score 28; DB 15; Length 734;  
Best Local Similarity 66.7%; Pred. No. 9.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LKXSLPERV 9  
Db 160 LAOSLPERL 168

RESULT 9  
US-10-369-493-18631  
; Sequence 18631, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 18631  
; LENGTH: 821  
; TYPE: PRT  
; ORGANISM: Halobacterium sp. NRC-1  
US-10-369-493-18631

Query Match 82.4%; Score 28; DB 15; Length 821;  
Best Local Similarity 66.7%; Pred. No. 1e+03;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LKXSLPERV 9  
Db 620 LKXSLPERL 628

RESULT 10  
US-10-369-493-2087  
; Sequence 2087, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 2087  
LENGTH: 925  
TYPE: PRT  
ORGANISM: Schizosaccharomyces pombe  
US-10-369-493-2087

Query Match 82.4%; Score 28; DB 15; Length 925;  
Best Local Similarity 66.7%; Pred. No. 1.2e+03;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLPERV 9  
DB 608 LSQSLPERL 616

RESULT 11  
US-10-408-765A-1706  
Sequence 1706, Application US/10408765A  
Publication No. US20040101874A1  
GENERAL INFORMATION:  
APPLICANT: Ghosh, Soumitra S.  
APPLICANT: Zhang, Bing  
APPLICANT: Gibson, Bradford W.  
APPLICANT: Taylor, Steven W.  
APPLICANT: Glenn, Gary M.  
APPLICANT: Matlock, Dale B.  
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
FILE REFERENCE: 660088.465  
CURRENT APPLICATION NUMBER: US/10/408,765A  
CURRENT FILING DATE: 2003-04-04  
NUMBER OF SEQ ID NOS: 3077  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1706  
LENGTH: 935  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-408-765A-1706

Query Match 82.4%; Score 28; DB 16; Length 935;  
Best Local Similarity 66.7%; Pred. No. 1.2e+03;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLPERV 9  
DB 680 LSQSLPERL 688

RESULT 12  
US-10-474-553-6  
Sequence 6, Application US/10474553  
Publication No. US20040161765A1  
GENERAL INFORMATION:  
APPLICANT: JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING DISEASE  
FILE REFERENCE: DRY-020.25  
CURRENT APPLICATION NUMBER: US/10/474,553  
CURRENT FILING DATE: 2003-10-10  
PRIOR APPLICATION NUMBER: 60/283,920

PRIOR FILING DATE: 2001-04-13  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 1118  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-474-553-6

Query Match 82.4%; Score 28; DB 16; Length 1118;  
Best Local Similarity 66.7%; Pred. No. 1.4e+03;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLPERV 9  
DB 680 LSQSLPERL 688

RESULT 13  
US-10-282-122A-67256  
Sequence 67256, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zykkind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PAM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 67256  
LENGTH: 210  
TYPE: PRT  
ORGANISM: Pasteurella multocida  
US-10-282-122A-67256

Query Match 79.4%; Score 27; DB 15; Length 210;  
Best Local Similarity 75.0%; Pred. No. 4.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLPER 8

Db 13 LARSLEPER 20

```

RESULT 14
US-10-437-963-142655
; Sequence 142655, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142655
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: clone ID: PAT_MRT4530_43640C.1.rep
US-10-437-963-142655

```

```

Query Match 79.4%; Score 27; DB 16; Length 286;
Best Local Similarity 55.6%; Pred. No. 6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 LXXSLPERV 9
Db 254 IASALPERV 262

```

```

RESULT 15
US-10-282-122A-68301
; Sequence 68301, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23

```

```

; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68301
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-68301

```

```

Query Match 79.4%; Score 27; DB 15; Length 298;
Best Local Similarity 55.6%; Pred. No. 6.3e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 LXXSLPERV 9
Db 49 LSHSLFORI 57

```

Search completed: April 18, 2005, 09:04:06  
Job time : 31.927 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 07:32:05 ; Search time 8.67153 Seconds  
(without alignments)  
99.861 Million cell updates/sec

Title: US-10-652-334-5  
Sequence: 1 LXXSLPERV 9

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: p1r1:.\*  
2: p1r2:.\*  
3: p1r3:.\*  
4: p1r4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	31	91.2	1121 2	S30862 DNA dependent ATPase
2	28	82.4	230 2	H83606 probable permease
3	28	82.4	341 2	AC2819 cyclopropane-fatty
4	28	82.4	341 2	D97597 cyclopropane fatty
5	28	82.4	451 2	P83793 DNA-damage-inducib
6	28	82.4	644 2	H64675 virulence associat
7	28	82.4	821 2	C84304 DNA helicase (mpo
8	28	82.4	935 2	S62476 hypothetical prote
9	28	82.4	1069 2	T33280 nonsense-mediated
10	28	82.4	1240 2	T06404 resistance complex
11	27	79.4	223 2	S76924 hypothetical prote
12	27	79.4	224 2	C83777 hypothetical prote
13	27	79.4	296 2	B36864 OAC2 protein_Azo
14	27	79.4	312 2	B88492 protein T0783.3 (i
15	27	79.4	333 2	H83219 probable pyruvate
16	27	79.4	361 2	F97068 histidinol-phospha
17	27	79.4	399 2	D95279 probable alcohol
18	27	79.4	411 2	D88504 protein B0361.7 (i
19	27	79.4	418 2	C84565 hypothetical prote
20	27	79.4	431 2	G82430 conserved hypotnet
21	27	79.4	485 2	H72424 hypothetical prote
22	27	79.4	495 1	S77406 protein kinase pkn
23	27	79.4	497 2	D83628 probable aldehyde
24	27	79.4	586 2	T24835 hypothetical prote
25	27	79.4	644 2	H71840 ribonuclease II fa
26	27	79.4	696 2	H83024 hypothetical prote
27	27	79.4	971 1	A36376 replication licens
28	27	79.4	971 2	S23408 prematurely termin
29	27	79.4	1075 2	C96682 protein F1E22.16 (

30	27	79.4	1118 2	S75309 hypothetical prote
31	26	76.5	249 2	B71024 probable lipote p
32	26	76.5	261 2	H70100 conserved hypotnet
33	26	76.5	265 2	AE3339 hypothetical prote
34	26	76.5	284 2	A83290 probable transcrip
35	26	76.5	304 2	JCS845 chitinase (EC 3.2.
36	26	76.5	309 2	T16660 hypothetical prote
37	26	76.5	310 2	G82279 transcription regu
38	26	76.5	324 2	T52609 cysteine synthase
39	26	76.5	333 2	D95283 probable Arac-type
40	26	76.5	383 2	H71848 probable na+/h+ an
41	26	76.5	386 2	A64114 probable ATPase m
42	26	76.5	391 2	T32706 hypothetical prote
43	26	76.5	433 2	H90495 metabolite transp
44	26	76.5	435 2	B86149 T1N6.17 protein -
45	26	76.5	640 2	A64065 hypothetical prote

## ALIGNMENTS

RESULT 1  
S30862  
DNA dependent ATPase/DNA helicase B - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein YER176w  
C:Species: Saccharomyces cerevisiae  
C:Date: 28-May-1993 #sequence revision 28-May-1993 #text\_change 09-Jul-2004  
C:Accession: S30862; S50679; J02490; PC2368  
R:Muligan, J.T.; Dietrich, P.S.; Henessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor, submitted to the EMBL Data Library, February 1993  
A:Reference number: S30812  
A:Accession: S30862  
A:Molecule type: DNA  
A:Residues: 1-1121 <MULT>  
A:Cross-references: UNIPROT:P32644; GB:U18922; EMBL:L11229; NID:9603405; PIDN:AA864703.1  
R:Dietrich, P.S.  
Submitted to the EMBL Data Library, December 1994  
A:Description: The sequence of S. cerevisiae cosmid 9163 and 9132.  
A:Reference number: S50679  
A:Accession: S50679  
A:Molecule type: DNA  
A:Residues: 1-1121 <DIB>  
A:Cross-references: EMBL:U18922; NID:9603405; PIDN:AA864703.1; PID:9603417; MIPS:YER176w  
R:Blawas, S.B.; Chen, P.H.; Leszyk, J.; Blawas, S.B.  
Biochem. Biophys. Res. Commun. 206, 850-856, 1995  
A:Title: Biochemical and genetic characterization of a replication protein A dependent D  
A:Reference number: J02490; MUID:95134267; PMID:7832796  
A:Molecule type: DNA  
A:Residues: 1-1121 <BIS>  
A:Accession: PC2368  
A:Molecule type: protein  
A:Residues: 277-283;623-633; 'X', 635-643 <B2>  
C:Comment: This enzyme plays pivotal roles in the unwinding of the DNA double helix dur  
C:Genetics:  
A:Gene: SGD:ECM32  
A:Cross-references: SGD:S0000978; MIPS:YER176w  
A:Map position: ER  
C:Keywords: nucleotide binding; P-loop  
P:670-677/Region: nucleotide-binding motif A (P-loop)  
Query Match 91.2%; Score 31; DB 2; Length 1121;  
Best Local Similarity 77.8%; Pred. No. 50;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 LXXSLPERV 9  
Db 860 LETSLPERV 868  
RESULT 2  
H83606  
probable permease of ABC transporter PA0313 (Imported) - Pseudomonas aeruginosa (strain

```

C:/Species: Pseudomonas aeruginosa
C:/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:/Accession: H83606
R:/Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:/Reference number: AB2950; MUID:20437337; PMID:10984043
A:/Accession: H83606
A:/Status: preliminary
A:/Molecule type: DNA
A:/Residues: 1-230 <STO>
A:/Cross-references: UNIPROT:Q916H8; GB:AE004469; GB:AE004091; NID:g9946152; PIDN:AG0370
C:/Genetics:
A:/Experimental source: strain PA01
A:/Gene: PA0313
C:/Superfamily: histidine permease protein M

Query Match 82.4%; Score 28; DB 2; Length 230;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Cy 1 LXSLEPRV 9
Db 208 LLSALFERY 216

RESULT 3
AC2819
Cyclopropane fatty acyl-phospholipid synthase [imported] - Agrobacterium tumefaciens (str
C:/Species: Agrobacterium tumefaciens
C:/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:/Accession: AC2819
R:/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monk, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, M.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClell
J.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Kreespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:/Reference number: AB2577; MUID:21608550; PMID:11743193
A:/Accession: AC2819
A:/Status: preliminary
A:/Molecule type: DNA
A:/Residues: 1-341 <KOR>
A:/Cross-references: UNIPROT:Q8UDZ2; GB:AE008688; PIDN:AAL42969.1; PID:G17740429; GSPDB:G
C:/Genetics:
A:/Experimental source: strain C58 (Dupont)
A:/Gene: cfa
A:/Map position: circular chromosome

Query Match 82.4%; Score 28; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 4 SLPERV 9
Db 201 SLFERY 206

RESULT 4
D97597
Cyclopropane fatty acyl-phospholipid synthase (AE005389) [imported] - Agrobacterium tumef
C:/Species: Agrobacterium tumefaciens
C:/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:/Accession: D97597
R:/Goodner, B.; Hinkley, G.; Gatlung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A:/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:/Reference number: AB7359; MUID:21608551; PMID:11743194
A:/Accession: D97597

```

```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-341 <KUR>
A:Cross-references: UNIPROT:O8UDZ2; GB:AE007869; PIDD:AAK87733.1; PID:G15157098; GSPDB:GT
A:Gene: AGR_C_3595
A:Map position: circular chromosome

Query Match
Best Local Similarity 100.0%; Score 28; DB 2; Length 341;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLFERV 9
|||
201 SLFERV 206

RESULT 5
F83793
DNA-damage-inducible protein BH1150 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #ext_change 09-Jul-2004
C:Accession: F83793
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: F83793
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-451 <STO>
A:Cross-references: UNIPROT:O9KDR1; GB:AP001511; GB:BA000004; NID:G10173727; PIDD:BA0486
A:Experimental source: strain C-125
A:Genetics:
A:Gene: BH1150

Query Match
Best Local Similarity 100.0%; Score 28; DB 2; Length 451;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLFERV 9
|||
DB 311 SLFERV 316

RESULT 6
H64675
virulence associated protein homolog - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #ext_change 09-Jul-2004
C:Accession: H64675
R:Tomson, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C.J.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: H64675
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-644 <TOM>
A:Cross-references: UNIPROT:D56123; GB:AE000630; GB:AE000511; NID:G2314409; PIDD:AA00829
A:Superfamily: virulence-associated protein vacB homolog

Query Match
Best Local Similarity 82.4%; Score 28; DB 2; Length 644;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLFERV 9
|||
DB 178 LXXSLFERV 186

```

RESULT 7  
DNA hellicase [imported] - Halobacterium sp. NRC-1  
C/Species: Halobacterium sp. NRC-1  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: C84304  
R/Name: W.V. Kennedy, S.P.; Mahltas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danesh, M.J.; Hough, D.W.; Madocke, D.G.; Jaldic  
Jung, K.H.; Alam, M.; Pretas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A/Title: Genome sequence of Halobacterium species NRC-1.  
A/Reference number: A84160; MUID:20504483; PMID:11016950  
A/Accession: C84304  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-821 <STO>  
A/Cross-references: UNIPROT:Q9HPR8; GB:AE004437; NID:g10580995; PIDN:AA61979.1; GSPDB:C  
C/Genetics:  
A/Gene: hel

Query Match 82.4%; Score 28; DB 2; Length 821;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLFRRV 9  
DB 620 LSKSLFRL 628

RESULT 8  
S62476  
hypothetical protein SPAC16C9.06c - fission yeast (Schizosaccharomyces pombe)  
C/Species: Schizosaccharomyces pombe  
C/Date: 16-May-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C/Accession: S62476; T37779  
R/Bedcock, K.; Churcher, C.M.  
submitted to the EMBL Data Library, October 1995  
A/Reference number: S62445  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-935 <BAD>  
A/Cross-references: UNIPROT:Q09820; EMBL:Z54366; NID:g1019812; PIDN:CAA91194.1; PID:g132  
R/Bedcock, K.; Churcher, C.M.; Barrett, B.G.; Raftery, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, October 1995  
A/Reference number: 221745  
A/Accession: T37779  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 11-935 <BA2>  
A/Cross-references: EMBL:Z54366; PIDN:CAA91194.2; GSPDB:GN00066; SPDB:SPAC16C9.06c  
C/Genetics:  
A/Gene: SPDB:SPAC16C9.06c  
A/Map position: 1L

Query Match 82.4%; Score 28; DB 2; Length 935;  
Best Local Similarity 66.7%; Pred. No. 2.1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLFRRV 9  
DB 618 LSGSLFRL 626

RESULT 9  
T43280  
nonsense-mediated mRNA decay trans-acting factor - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C/Accession: T43280

R/Page, M.F.; Carr, B.; Anderg, K.R.; Grimson, A.; Anderson, P.  
Mol. Cell. Biol. 19, 5943-5951, 1999  
A/Title: SMG-2 is a phosphorylated protein required for mRNA surveillance in Caenorhabditis  
A/Reference number: Z22389; MUID:99384262; PMID:10454541  
A/Accession: T43280  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-1069 <PAG>  
A/Cross-references: UNIPROT:O76512; EMBL:AF074017; NID:g3328176; PIDN:AA626789.1; PID:g33

Query Match 82.4%; Score 28; DB 2; Length 1069;  
Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLFRRV 9  
DB 654 LSGSLFRL 662

RESULT 10  
T06404  
resistance complex protein I2C-2 - tomato  
C/Species: Lycopersicon esculentum (tomato)  
C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C/Accession: T06404  
R/Ori, N.; Eshed, Y.; Paran, I.; Presting, G.; Aviv, D.; Tankale, S.; Zamir, D.; Fluhr,  
Plant Cell 9, 521-532, 1997  
A/Title: The I2C family from the wilt disease resistance locus I2 belongs to the nucleot  
A/Reference number: Z15652; MUID:97290204; PMID:9144960  
A/Accession: T06404  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1240 <ORI>  
A/Cross-references: UNIPROT:O24016; EMBL:AF004879; NID:g2258316; PIDN:AA63275.1; PID:g2  
C/Genetics:  
A/Gene: I2C-2  
A/Map position: 11  
A/Function:  
A/Description: confers resistance against Fusarium oxysporum  
C/Superfamily: disease resistance protein RPS2; leucine-rich alpha-2-glycoprotein repeat

Query Match 82.4%; Score 28; DB 2; Length 1240;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLFRRV 9  
DB 472 SLFRRV 477

RESULT 11  
S76924  
hypothetical protein - Synechocystis sp. (strain PCC 6803)  
C/Species: Synechocystis sp.  
A/Variety: PCC 6803  
C/Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C/Accession: S76924  
R/Kaneko, T.; Sato, S.; Kocant, H.; Tanaka, A.; Asami, B.; Nakamura, Y.; Miyajima, N.;  
O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-116, 1996  
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis  
S.  
A/Reference number: S74322; MUID:97061201; PMID:8905231  
A/Accession: S76924  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-223 <KAN>  
A/Cross-references: UNIPROT:P74717; EMBL:D90917; GB:AE001339; NID:g1653836; PIDN:BA1883  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 79.4%; Score 27; DB 2; Length 223;  
Best Local Similarity 83.3%; Pred. No. 85;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLFPERV 9  
|||||:  
Db 166 SLFPERI 171

## RESULT 12

hypothetical protein BH1019 [imported] - Bacillus halodurans (strain C-125)  
C/Species: Bacillus halodurans  
C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C/Accession: C83777  
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A/Reference number: A83650; MUID:20512582; PMID:11058132  
A/Accession: C83777  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-224 <STO>  
A/Cross-references: UNIPROT:Q9KE39; GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BAH047  
A/Experimental source: strain C-125  
C/Genetics:  
A/Gene: BH1019  
C/Superfamily: Bacillus subtilis hypothetical protein yhzB

Query Match 79.4%; Score 27; DB 2; Length 224;  
Best Local Similarity 55.6%; Pred. No. 86;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLFPERV 9  
: |||||  
Db 6 IESSIFERV 14

## RESULT 13

B36864  
OAC2 protein - Azorhizobium caulinodans  
C/Species: Azorhizobium caulinodans  
C/Date: 02-Jun-1995 #sequence\_revision 02-Jun-1995 #text\_change 09-Jul-2004  
C/Accession: B36864; S32661  
R/Goethals, K.; Leyman, B.; Van Den Bede, G.; Van Montagu, M.; Holsters, M.  
J. Bacteriol. 176, 92-99, 1994  
A/Title: An Azorhizobium caulinodans ORS571 locus involved in lipopolysaccharide product  
A/Reference number: A36864; MUID:94110246; PMID:7506708  
A/Accession: B36864  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-296 <GOB>  
A/Cross-references: UNIPROT:Q06936; EMBL:Z22611; NID:g296706; PIDN:CAA80331.1; PID:g2967  
C/Superfamily: dmp-dihydrostreptose synthase

Query Match 79.4%; Score 27; DB 2; Length 296;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLFPER 8  
: |||||  
Db 260 LDSSLFPER 267

## RESULT 14

B88492  
proteins T07E3.3 [imported] - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C/Accession: B88492  
R:anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A/Reference number: A75000; MUID:99069613; PMID:9851916  
A/Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A/Accession: B88492  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-312 <STO>  
C/Genetics:  
A/Gene: T07E3.3  
A/Map position: 3

Query Match 79.4%; Score 27; DB 2; Length 312;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLFPERV 9  
|||||:  
Db 30 SLFPERI 35

## RESULT 15

H83219  
probable pyruvate dehydrogenase E1 component, beta chain PA3416 [imported] - Pseudomonas  
C/Species: Pseudomonas aeruginosa  
C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 12-Jul-2004  
C/Accession: H83219  
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
Loiy, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A/Reference number: A82950; MUID:20437337; PMID:10984043  
A/Accession: H83219  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-333 <STO>  
A/Cross-references: UNIPROT:Q9HY19; GB:AE004762; GB:AE004091; NID:g949544; PIDN:AAG0680  
A/Experimental source: strain PA01  
C/Genetics:  
A/Gene: PA3416  
C/Superfamily: pyruvate dehydrogenase, E1 component, beta subunit

Query Match 79.4%; Score 27; DB 2; Length 333;  
Best Local Similarity 55.6%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXXSLFPERV 9  
: |||||  
Db 280 IMASLFPERV 288

Search completed: April 18, 2005, 08:06:03  
Job time : 10.6715 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 18, 2005, 07:33:55 ; Search time 37.5766 Seconds  
(without alignments)

122.648 Million cell updates/sec

Title: US-10-652-334-5  
Perfect score: 34  
Sequence: 1 LXXSLFERV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	91.2	1121	1	YE06 YEAST
2	31	91.2	1124	2	Q6BNH2
3	30	88.2	378	2	Q6B079
4	30	88.2	867	2	Q6CDN7
5	30	88.2	1125	2	Q6EKF3
6	29	85.3	378	2	Q6PRQ1
7	29	85.3	378	2	Q6PSN2
8	29	85.3	1087	2	Q7YYP3
9	28	82.4	47	2	Q7MH8
10	28	82.4	201	1	CBIT_THERVO
11	28	82.4	230	2	Q916H8
12	28	82.4	287	2	Q6R7F6
13	28	82.4	341	2	Q8UD22
14	28	82.4	365	2	Q7VJPO
15	28	82.4	422	2	Q6SUP7
16	28	82.4	451	2	Q9KDL1
17	28	82.4	478	2	Q8PG59
18	28	82.4	502	2	Q6EVG7
19	28	82.4	502	2	Q6EVH3
20	28	82.4	512	2	Q7L804
21	28	82.4	518	2	Q8JME0
22	28	82.4	532	2	Q9VCQ0
23	28	82.4	533	2	Q9Y2F0
24	28	82.4	543	2	Q8KON4
25	28	82.4	598	2	Q9BML1
26	28	82.4	619	2	Q8RL7
27	28	82.4	636	2	Q6BKG5
28	28	82.4	644	1	KNR_HELPY
29	28	82.4	764	1	Q6F2C1
30	28	82.4	821	2	Q9HPR8
31	28	82.4	925	1	RNT1_SCHPO

32	28	82.4	1069	1	RNT1_CAERL	Q76512 caenorhabdi
33	28	82.4	1097	1	RNT1_FUGRU	Q96123 fugu rubrip
34	28	82.4	1098	2	Q6GNR2	Q6GNR2 xenopus lae
35	28	82.4	1100	2	Q7VZV4	Q7VZV4 brachydanio
36	28	82.4	1113	1	RNT1_MOUSE	Q96P00 mus musculu
37	28	82.4	1113	2	Q6GIF5	Q6GIF5 mus musculu
38	28	82.4	1118	2	Q66Z25	Q66Z25 homo sapien
39	28	82.4	1120	2	Q7PWZ4	Q7PWZ4 anopheles g
40	28	82.4	1124	2	Q6PHQ5	Q6PHQ5 mus musculu
41	28	82.4	1129	1	RNT1_HUMAN	Q92900 homo sapien
42	28	82.4	1180	1	RNT1_DROME	Q9YJ63 drosophila
43	28	82.4	1235	1	RNT1_ARATH	Q9FJ70 arabidopsis
44	28	82.4	1240	2	Q24016	Q24016 lycopersico
45	28	82.4	1243	2	Q6S3K7	Q6S3K7 arabidopsis

## ALIGNMENTS

RESULT 1  
YE06 YEAST STANDARD, PRT, 1121 AA.  
ID YE06 YEAST  
AC P32644  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Hypothetical 127.0 kDa protein in RAD24-BM1 intergenic region.  
GN OrderedLocustNames=VER176W; ORFNames=SYCP-ORF61;  
OS Saccharomyces cerevisiae (baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RS SEQUENCE FROM N.A.  
RC STRAIN=S288c / AB972;  
RX MEDLINE=97313264; PubMed=916968;  
RA Arai R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,  
RA Cherry J.M., Chung E., Duncan M., Guzman T., Hartzell G.,  
RA Huntke-Smith S., Hyman R.W., Kayser A., Komp C., Laehkari D., Lew H.,  
RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,  
RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,  
RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;  
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";  
RL Nature 387:78-81(1997).  
CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcement/>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC -----  
DR EMBL: U18922; AAB64703.1; --  
DR PIR: S30862; S30862.  
DR GERMOnline: 139253; --  
DR SGD: S00000978; ECM32.  
DR GO: GO:0005844; C:polyome; IDA.  
DR GO: GO:0003678; F:DNA helicase activity; IDA.  
DR GO: GO:0006449; P:regulation of translational termination; IMP.  
DR InterPro: IPR001410; DRAD.  
DR SMART: SM00487; DEXDC; 1.  
KW ATP-binding, Helicase, Hypothetical protein.  
KW NP BIND 670 677 ATP (Potential).  
SQ SEQUENCE 1121 AA; 126970 MW; 641C4AA6810282A0 CRC64;  
Query Match 91.2%; Score 31; DB 1; Length 1121;  
Best Local Similarity 77.8%; Pred. No. 2e+02; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 2;  
1 LXXSLFERV 9

Db 860 LEXSLFERNV 868

## RESULT 2

ID Q6BNH2 PRELIMINARY; PRT; 1124 AA.

AC Q6BNH2; 25-OCT-2004 (TRENBLrel. 28, Created)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)  
 DE Debaryomyces hansenii chromosome B of strain CBS767 of Debaryomyces hansenii.  
 GN ORFNames=DEHA0E23034g;  
 OS Debaryomyces hansenii CBS767.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Debaryomyces.  
 CX NCBI\_TaxID=284592;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS767;  
 RG Genolevures;  
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marck C., Nevegilise C., Talla E.,  
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Battey S., Blanchin S., Beckerich J.M., Beyne E., Bleykaeten C.,  
 RA Boissarie A., Boyer J., Catolico L., Confanioleri F., de Daruvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
 RA Hantreay F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicaud J.M., Nikoliski M., Ozias S., Ozler-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
 RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
 RA Zelenko-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,  
 RA Wincker P., Souciet J.L.;  
 RT "Genome evolution in yeasts."  
 RL Nature 430:35-44(2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS767;  
 RA Genoscope;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR362137; CAG88524.1;  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR InterPro; IPR003593; AAA\_Arpase.  
 DR SMART; SM00382; AAA; 1.  
 DR PROSITE; PS00659; GLYCOSYL\_HYDROL\_F5; UNKNOWN\_1.  
 DR ATP-binding.  
 SQ SEQUENCE 1124 AA; 127842 MW; D6F8B0958614EE9 CRC64;  
 Query Match 91.2%; Score 31; DB 2; Length 1124;  
 Best Local Similarity 77.8%; Pred. No. 2e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

## RESULT 3

ID Q80079 PRELIMINARY; PRT; 378 AA.

AC Q80079; 01-OCT-2002 (TRENBLrel. 22, Created)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)  
 DE Transposase.  
 GN OrderedLocusNames=MM0258;  
 OS Methanosarcina mazei (Methanosarcina frisia).  
 CC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;  
 OC Methanosarcinaceae; Methanosarcina.

OX NCBI\_TaxID=2209;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Goel / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;  
 RX MEDLINE=22120827; PubMed=12125824;  
 RA Deppemeier U., Johann A., Hartach T., Merl R., Schmitz R.A.,  
 RA Martinez-Arias R., Henne A., Wierer A., Baumeister S., Jacobi C.,  
 RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,  
 RA Bhattacharyya A., Lykilds A., Overbeek R., Klenk H.-P., Gunsalus R.P.,  
 RA Filtz H.-J., Gottschalk G.;  
 RT "The genome of Methanosarcina mazei: evidence for lateral gene transfer between Bacteria and Archaea."  
 RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIB99;  
 RA Genoscope;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE013250; AA29954.1;  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0004803; F:transposase activity; IEA.  
 DR GO; GO:0006313; P:DNA transposition; IEA.  
 DR InterPro; IPR002559; Transposase\_11.  
 DR Pfam; PF01609; Transposase\_11; 1.  
 RN [2]  
 RP Complete proteome.  
 SQ SEQUENCE 378 AA; 44016 MW; 0C40F54EF713CB60 CRC64;

Query Match 88.2%; Score 30; DB 2; Length 378;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LEXSLFERNV 9

Db 159 LEXSLFERNV 167

## RESULT 4

ID Q6CDN7 PRELIMINARY; PRT; 867 AA.

AC Q6CDN7; 25-OCT-2004 (TRENBLrel. 28, Created)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)  
 DE Yarrowia lipolytica chromosome B of strain CLIB99 of Yarrowia lipolytica.  
 GN ORFNames=YALI0B2528g;  
 OS Yarrowia lipolytica CLIB99.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Dipodascaceae; Yarrowia.  
 CX NCBI\_TaxID=284591;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIB99;  
 RG Genolevures;  
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marck C., Nevegilise C., Talla E.,  
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Battey S., Blanchin S., Beckerich J.M., Beyne E., Bleykaeten C.,  
 RA Boissarie A., Boyer J., Catolico L., Confanioleri F., de Daruvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
 RA Hantreay F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicaud J.M., Nikoliski M., Ozias S., Ozler-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
 RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
 RA Zelenko-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,  
 RA Wincker P., Souciet J.L.;  
 RT "Genome evolution in yeasts."  
 RL Nature 430:35-44(2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIB99;  
 RA Genoscope;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR362138; CAG83478.1;  
 DR GO; GO:0005524; F:ATP binding; IEA.

DR GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase like.  
 DR InterPro: IPR007719; Prot. Kinase.  
 DR InterPro: IPR02290; Ser. Thr. Kinase.  
 DR InterPro: IPR08271; Ser. Thr. Pkin. AS.  
 DR InterPro: IPR01245; Tyr. Pkinase.  
 DR Pfam: PF00069; Pkinase\_1.  
 DR ProDom: PD000001; Prot. Kinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR SMART: SM00219; TyKc; 1.  
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE: PS0011; PROTEIN KINASE DOM; 1.  
 DR PROSITE: PS00108; PROTEIN KINASE ST; 1.  
 KM ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 867 AA; 96798 MW; F86DC7819B99A1D CRC64;

Query Match 88.2%; Score 30; DB 2; Length 867;  
 Best Local Similarity 66.7%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LXXSLPERV 9  
 Db 454 LXXSLPERV 462

RESULT 5  
 Q6PFK3 PRELIMINARY; PRT; 1125 AA.  
 AC Q6PFK3  
 DT 05-JUN-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Candida glabrata strain CBS138 chromosome I complete sequence.  
 GN ORFNames=CAGI01120349;  
 OS Candida glabrata CBS138.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 NCBI\_TaxID=284593;  
 OX NCBI\_TaxID=284593;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS138;  
 RG Genolevures;  
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,  
 RA Goffard N., Frangeul L., Aigle M., Anthonard V., Babour A., Barde V.,  
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boismay A., Boyer J., Catolico L., Confanioli F., de Daruvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
 RA Hantziye F., Hennequin C., Jauniaux N., Joyet F., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Niclaud J.M., Nikolajski M., Oztas S., Ozler-Kalogeropoulos O.,  
 RA Pellenz S., Potler S., Richard G.F., Straub M.L., Suleau A.,  
 RA Szwennne D., Tekala F., Wesolowski-Louvel M., Westhof B., Wirth B.,  
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Pukhbar M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Galliard C., Weissenbach J.,  
 RA Wincker P., Souciet J.L.;  
 RT "Genome evolution in yeasts";  
 RL Nature 430:35-44(2004).  
 DR EMBL: CR380958; CAG62265.1;  
 SQ SEQUENCE 1125 AA; 127137 MW; 63C0428123F3CC8C CRC64;

Query Match 88.2%; Score 30; DB 2; Length 1125;  
 Best Local Similarity 66.7%; Pred. No. 3.5e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LXXSLPERV 9  
 Db 865 LXXSLPERV 873

RESULT 6

Q8PRO1 PRELIMINARY; PRT; 378 AA.  
 AC Q8PRO1  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Transposase.  
 GN OrderedLocustNames=MM2507, MM2509, MM3265;  
 OS Methanosarcina mazel (Methanosarcina fistula).  
 OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;  
 OC Methanosarcinaceae; Methanosarcina.  
 NCBI\_TaxID=2209;  
 OX NCBI\_TaxID=2209;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Goel / ATCC BAA-199 / DSM 3647 / OCM 88;  
 RX MEDLINE=22120827; PubMed=12125824;  
 RA Deppe-meier U., Johann A., Hartesch T., Merl R., Schmitz R.A.,  
 RA Martinez-Arias R., Henne A., Mieser A., Baumer S., Jacobi C.,  
 RA Bruggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,  
 RA Bhattacharya A., Lykilds A., Overbeek R., Klenk H.-P., Gunsalus R.P.,  
 RA Fritz H.-J., Gottschalk G.;  
 RT "The genome of Methanosarcina mazel: evidence for lateral gene  
 transfer between Bacteria and Archaea.";  
 RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).  
 DR EMBL: AE013494; AAM32203.1;  
 DR EMBL: AE013585; AAM32961.1;  
 DR EMBL: AE013495; AAM32205.1;  
 DR GO:0003677; F:DNA binding; IEA.  
 DR GO:0004803; F:transposase activity; IEA.  
 DR GO:0006313; P:DNA transposition; IEA.  
 DR InterPro: IPR02559; Transposase\_11.  
 DR Pfam: PF01609; Transposase\_11; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 378 AA; 43945 MW; C82EBA65F59A0B70 CRC64;

Query Match 85.3%; Score 29; DB 2; Length 378;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LXXSLPERV 9  
 Db 159 LXXSLPERV 167

RESULT 7  
 Q8PSN2 PRELIMINARY; PRT; 378 AA.  
 AC Q8PSN2  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Transposase.  
 GN OrderedLocustNames=MM3047;  
 OS Methanosarcina mazel (Methanosarcina fistula).  
 OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;  
 OC Methanosarcinaceae; Methanosarcina.  
 NCBI\_TaxID=2209;  
 OX NCBI\_TaxID=2209;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Goel / ATCC BAA-199 / DSM 3647 / OCM 88;  
 RX MEDLINE=22120827; PubMed=12125824;  
 RA Deppe-meier U., Johann A., Hartesch T., Merl R., Schmitz R.A.,  
 RA Martinez-Arias R., Henne A., Mieser A., Baumer S., Jacobi C.,  
 RA Bruggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,  
 RA Bhattacharya A., Lykilds A., Overbeek R., Klenk H.-P., Gunsalus R.P.,  
 RA Fritz H.-J., Gottschalk G.;  
 RT "The genome of Methanosarcina mazel: evidence for lateral gene  
 transfer between Bacteria and Archaea.";  
 RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).  
 DR EMBL: AE013560; AAM32743.1;  
 DR GO:0003677; F:DNA binding; IEA.  
 DR GO:0004803; F:transposase activity; IEA.  
 DR GO:0006313; P:DNA transposition; IEA.

DR InterPro: IPR002559; Transposase\_11.  
 DR Pfam: PF01609; Transposase\_11; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 378 AA; 43944 MW; FE2B7C6DD59A0B78 CRC64;

Query Match 85.3%; Score 29; DB 2; Length 378;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LKXSLFERV 9  
 Db 159 LMSLPERI 167

RESULT 8  
 ID Q7YYP3 PRELIMINARY; PRT; 1087 AA.

AC Q7YYP3; 01-OCT-2003 (TRENBLrel. 25, Created)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE Elongation factor-like protein.  
 GN ORNames=IMB-551;  
 OS Cryptosporidium parvum.  
 OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;  
 OC Cryptosporidiidae; Cryptosporidium.  
 OX NCBI\_TaxID=5807;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bankier A.T., Spriggs H.F., Fartmann B., Konfortov B.A., Madera M.,  
 RA Vogel C., Teichmann S.A., Ivens A., Dear P.H.;  
 RT "Integrated mapping, chromosomal sequencing and sequence analysis of  
 RT Cryptosporidium parvum";  
 RL Genome Res. 0:0-0(2003).  
 DR EMBL; BX58352; CAD98437.1; -  
 DR GO; GO:0005525; F:GTP binding; IEA.  
 DR GO; GO:0003746; F:translation elongation factor activity; IEA.  
 DR GO; GO:0006412; P:protein biosynthesis; IEA.  
 DR GO; GO:0006414; P:translational elongation; IEA.  
 DR InterPro: IPR000640; EFG C.  
 DR InterPro: IPR009022; EFG-T11-V.  
 DR InterPro: IPR00517; EFG-IV.  
 DR InterPro: IPR000795; ProtSyn GTPbind.  
 DR InterPro: IPR005225; Small\_GTP.  
 DR Pfam; PF00679; EFG\_C; 1.  
 DR Pfam; PF03764; EFG\_IV; 1.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR TIGRFAMs; TIGR00231; small\_GTP; 1.  
 KW Elongation factor; GTP-binding; Protein biosynthesis.  
 SQ SEQUENCE 1087 AA; 123753 MW; 0E0158224886F83 CRC64;

Query Match 85.3%; Score 29; DB 2; Length 1087;  
 Best Local Similarity 66.7%; Pred. No. 5.9e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LKXSLFERV 9  
 Db 353 LKXSLFERV 361

RESULT 9  
 ID Q7MHH8 PRELIMINARY; PRT; 47 AA.

AC Q7MHH8; 01-MAR-2004 (TRENBLrel. 26, Created)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE Hypothetical protein VV2891.  
 GN OrderedLocusNames=VV2891;  
 OS Vibrio vulnificus (strain VJ016).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=196600;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=14656965; DOI=10.1101/9r.1295503;  
 RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,  
 RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,  
 RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;  
 RT "Comparative genome analysis of *Vibrio vulnificus*, a marine  
 RT pathogen";  
 RL Genome Res. 13:2577-2587(2003).  
 DR EMBL; AP005341; BAC95655.1; -  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 47 AA; 5454 MW; 45849EC67B07C8D3 CRC64;

Query Match 82.4%; Score 28; DB 2; Length 47;  
 Best Local Similarity 66.7%; Pred. No. 48;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LKXSLFERV 9  
 Db 12 LKXSLFERV 20

RESULT 10  
 ID CBIT\_THEVO STANDARD; PRT; 201 AA.

AC Q97A64; 05-JUL-2004 (Rel. 44, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Probable cobalt-precorrin-6Y C(15)-methyltransferase [decarboxylating]  
 DE (EC 2.1.1.-).  
 GN Name=cbtI; OrderedLocusNames=TV0946; ORFNames=TVG0972073;  
 GN Thermoplasma volcanum.  
 OS Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;  
 OC Thermoplasmataceae; Thermoplasmata.  
 OX NCBI\_TaxID=50339;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RP STRAIN=GS61 / DSM 4299 / JCM 9571;  
 RC MEDLINE=20570466; PubMed=11121031; DOI=10.1073/pnas.97.26.14257;  
 RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,  
 RA Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,  
 RA Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;  
 RT "Archaeal adaptation to higher temperatures revealed by genomic  
 RT sequence of *Thermoplasma volcanum*";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).

-1- FUNCTION: Probably catalyzes the methylation of either C-15 or C-5  
 in cobalt-precorrin-6Y to form cobalt-precorrin-7M. Methylation of  
 C-15 would probably be followed by a spontaneous decarboxylation  
 of C-12 (By similarity).

-1- PATHWAY: Adenosylcobalamin biosynthesis; anaerobic branch of  
 corrin ring synthesis; tenth step.  
 CC -1- SIMILARITY: Belongs to the methyltransferase superfamily. Archaeal-  
 type cbtI family.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; AP000994; BAB60088.1; -  
 DR HAMAP; MF\_00786; -; 1.  
 DR InterPro; IPR000051; SAM\_bind.  
 DR COBALTamin biosynthesis; Complete proteome; Methyltransferase;  
 KW Transferrase.  
 SQ SEQUENCE 201 AA; 21884 MW; 39D5832AA843A9A CRC64;

Query Match 82.4%; Score 28; DB 1; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLFERV 9  
17 SLFERV 22

## RESULT 11

OY16H8 PRELIMINARY; PRT; 230 AA.

AC 0916H8  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE Probable permease of ABC transporter.  
GN OrderedLocustNames=PA0313;  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=287;

## SEQUENCE FROM N.A.

RC STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman P.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Polger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.B.W., Lory S., Olson M.V.,  
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen.";  
RL Nature 406:959-964(2000).  
CC -1- FUNCTION: Part of a binding-protein-dependent transport system.  
CC Probably responsible for the translocation of the substrate across  
CC the membrane (By similarity).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
CC -1- SIMILARITY: Belongs to the binding-protein-dependent transport  
CC system permease family.  
EMBL, AE004469; AAC03702.1; -.  
DR PIR; H83606; H83606.  
DR GO; GO:0016021; C:Integral to membrane; IEA.  
DR GO; GO:0005215; P:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000515; BPD\_transp.  
DR InterPro; IPR010065; HEORC\_ABC\_3TM.  
DR Pfam; PF00528; BPD\_transp\_1; 1.  
DR TIGRfams; TIGR01726; HEORC\_pern\_3TM; 1.  
DR PROSITE; PS50928; ABC\_TM1; 1.  
KW Complete proteome; Transmembrane; Transport.  
SQ SEQUENCE 230 AA; 24963 MW; EB984ABE7968BA27 CRC64;

Query Match 82.4%; Score 28; DB 2; Length 230;  
Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
OY 1 LXXSLFERV 9  
DB 208 LLSALFERV 216

## RESULT 12

OY16H8 PRELIMINARY; PRT; 287 AA.

AC 06F7F6  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DE Hypothetical protein.  
GN OrderedLocustNames=ACIAD340;  
OS Acinetobacter sp. (strain ADP1).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Moraxellaceae; Acinetobacter.  
OX NCBI\_TaxID=62977;

RN [1]  
RP SEQUENCE FROM N.A.  
RA Barbe V., Vallonet D., Fonknechten N., Kreilmeyer A., Ozias S.,  
RA Labarre L., Cruevelier S., Robert C., Duprat S., Wincker P.,  
RA Ornston L.N., Weisenbach J., Marliere P., Cohen G.N., Medigue C.,  
RT ADP1, a versatile and naturally transformation competent bacterium.";  
RL Nucleic Acids Res. 0:0-0(2004).  
DR EMBL; CR543861; CAG70009.1; -.  
DR GO; GO:0016810; P:hydrolyase activity, acting on carbon-nitrog. . .; IEA.  
DR GO; GO:0006807; P:nitrogen metabolism; IEA.  
DR InterPro; IPR003010; NClse/CNhydase.  
DR InterPro; IPR001110; UPF0012.  
DR Pfam; PF00795; CN\_hydrolase; 1.  
DR PROSITE; PS50263; CN\_HYDROLASE; 1.  
DR PROSITE; PS01227; UPF0012; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 287 AA; 32191 MW; D4DCDE6B78106FB8 CRC64;

Query Match 82.4%; Score 28; DB 2; Length 287;  
Best Local Similarity 100.0%; Pred. No. 2.0e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLFERV 9  
DB 82 SLFERV 87

## RESULT 13

OY16H8 PRELIMINARY; PRT; 341 AA.

AC 08UD22, Q7CY60;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
DE Cyclopropane-fatty-acyl-phospholipid synthase (AGR\_C\_3595p).  
GN Name=cfa; OrderedLocustNames=AGR\_C\_3595; At11974;  
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
OX NCBI\_TaxID=176299;  
EMBL, AE008115; AAK87733.1; -.

DR PIR; H83606; H83606.  
DR GO; GO:0016021; C:Integral to membrane; IEA.  
DR GO; GO:0005215; P:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000515; BPD\_transp.  
DR InterPro; IPR010065; HEORC\_ABC\_3TM.  
DR Pfam; PF00528; BPD\_transp\_1; 1.  
DR TIGRfams; TIGR01726; HEORC\_pern\_3TM; 1.  
DR PROSITE; PS50928; ABC\_TM1; 1.  
KW Complete proteome; Transmembrane; Transport.  
SQ SEQUENCE 230 AA; 24963 MW; EB984ABE7968BA27 CRC64;

Query Match 82.4%; Score 28; DB 2; Length 287;  
Best Local Similarity 100.0%; Pred. No. 2.0e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LXXSLFERV 9  
DB 82 SLFERV 87

## SEQUENCE FROM N.A.

RC STRAIN=Cereon;  
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;  
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard C., Mullin L.,  
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
RA Houmel K., Gordon J., Vaudin M., Iatchouk O., Bp A., Liu F.,  
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,  
RA Cielo C., Slater S.;  
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
C58.";  
RL Science 294:2317-2323(2001).  
EMBL, AE008115; AAK87733.1; -.

```

DR PIR; AC2819; AC2819.
DR PIR; D97597; D97597.
DR GO; GO:000885; F:cytlopropane-fatty-acyl-phospholipid syntha. .; IEA.
DR GO; GO:000860; P:lipid biosynthesis; IEA.
DR InterPro; IPR003333; CMA5.
DR InterPro; IPR00276; GPCR_Rhodpsn.
DR InterPro; IPR001601; Methyltransf.
DR Pfam; PF03353; CMA5; 1.
DR PROSITE; PS00237; G_PROTEIN_RECER_FL_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 341 AA; 39042 MW; 7DAF3DBA2F142E4 CRC64;

Query Match 82.4%; Score 28; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLFPRV 9
DB 201 SLFPRV 206

RESULT 14
QYVJPO PRELIMINARY; PRT; 366 AA.
ID QYVJPO;
AC QYVJPO;
DT 01-OCT-2003 (TREMBlrel. 25; Created)
DT 01-OCT-2003 (TREMBlrel. 25; Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26; Last annotation update)
DE UDP-galactopyranose mutase (EC 5.4.99.9).
GN Name=glt; OrderedLocNames=HH0203;
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP STRAIN=ATCC 51449 / 3B1;
RX MEDLINE=22709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;
RA Suerbaum S., Josenhans C., Stenzenbach T., Drescher B., Brandt P.,
RA Bell M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koenig J., Macko L., Menz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox U.G.;
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus."
RT Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
DR EMBL; AE017144; AAP76800.1; -.
DR HSP; P37747; 118T.
DR GO; GO:0008767; F:UDP-galactopyranose mutase activity; IEA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro; IPR004379; UDP-GALP_mutase.
DR Pfam; PF03275; GLF; 1.
DR TIGRPFAMs; TIGR00031; UDP-GALP_mutase; 1.
KW Complete proteome.
SQ SEQUENCE 366 AA; 42921 MW; E013DF9BD4FC2BC8 CRC64;

Query Match 82.4%; Score 28; DB 2; Length 366;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LXXSLFPRV 9
DB 166 LXXSLFPRV 174

RESULT 15
QYVJPO PRELIMINARY; PRT; 422 AA.
ID QYVJPO;
AC QYVJPO;
DT 01-DEC-2001 (TREMBlrel. 19; Created)
DT 01-DEC-2001 (TREMBlrel. 19; Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24; Last annotation update)
DE Potassium inwardly rectifying channel DIR.

```

```

GN Name=Ir;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22126605; PubMed=12128223; DOI=10.1016/S0925-4773(02)00140-5;
RA Maclean S.J., Andrews B.C., Verheijen E.M.;
RT "Characterization of Dir: a putative potassium inward rectifying
RT channel in Drosophila."
RL Mech. Dev. 116:193-197(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the inward rectifier-type potassium channel
CC family.
DR EMBL; AY044166; AKG98804.1; -.
DR HSP; P35562; IN9P.
DR FLYBase; FBgn039061; Ir.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005242; F:Inward rectifier potassium channel activity; IEA.
DR GO; GO:0005267; F:Potassium channel activity; IEA.
DR GO; GO:0005244; F:Voltage-gated ion channel activity; IEA.
DR GO; GO:0006811; P:Ion transport; IEA.
DR GO; GO:0006813; P:Potassium ion transport; IEA.
DR InterPro; IPR01838; K-channel_IR.
DR InterPro; IPR01622; K-channel_pore.
DR Pfam; PF01007; IRK; 1.
DR PRINTS; PR01320; KIRCHANNEL.
KW Ion transport; Ionic channel; Potassium; Potassium transport;
KW Transmembrane; Transport; Voltage-gated channel.
SQ SEQUENCE 422 AA; 48457 MW; 40AC1956B9B29C3B CRC64;

Query Match 82.4%; Score 28; DB 2; Length 422;
Best Local Similarity 55.6%; Pred. No. 4.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LXXSLFPRV 9
DB 374 LXXSLFPRV 382

Search completed: April 18, 2005, 08:15:45
Job time : 39.5766 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using BW model

Run on: April 18, 2005, 07:10:15 ; Search time 104.54 Seconds  
(without alignments)  
77.692 Million cell updates/sec

Title: US-10-652-334-6  
Perfect score: 90  
Sequence: 1 LXXQYRMHPXISRPYXGXL 21

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004as:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	78	86.7	21	7	ABW01182 Saccharom
2	78	86.7	21	8	ADP44108 Yeast tra
3	78	86.7	22	3	AAV77809 Moticl VI
4	67.5	75.0	415	3	AAV77813 Yeast Mct
5	67.5	75.0	415	7	ABW01201 Saccharom
6	67.5	75.0	415	8	ADP44127 Yeast hel
7	65.5	72.8	925	8	ADN19434 Bacterial
8	65.5	72.8	992	8	ADN23757 Bacterial
9	64.5	71.7	471	3	AAV77815 Yeast Sen
10	64.5	71.7	472	7	ABW01202 Saccharom
11	64.5	71.7	472	8	ADP44128 Yeast hel
12	64.5	71.7	2000	6	ABR53451 Protein s
13	64.5	71.7	2000	7	ADK64622 Disease t
14	64.5	71.7	2231	8	ADN19177 Bacterial
15	61.5	68.3	771	8	ABO58160 Human gen
16	61.5	68.3	797	2	AAW36508 Human REN
17	61.5	68.3	935	7	ADJ69900 Human hea
18	61.5	68.3	1043	2	AAW36509 Murine RE
19	61.5	68.3	1118	6	ABG73900 Human REN
20	61.5	68.3	1140	2	AAV73777 Human HUP
21	59.5	66.1	380	3	AAV77814 Yeast Upf
22	59.5	66.1	380	7	ABW01205 Saccharom
23	59.5	66.1	380	8	ADP44131 Yeast hel
24	59.5	66.1	642	8	ADS43096 Bacterial
25	59.5	66.1	818	8	ADK68058 Female re

26	59.5	66.1	971	3	AAV98057	AAV98057 Yeast Upf
27	59.5	66.1	971	6	ABR53412	ABR53412 Protein s
28	59.5	66.1	971	7	ADK64706	ADK64706 Disease t
29	53.5	59.4	698	7	ADM25464	ADM25464 Hyperther
30	53	58.9	1944	8	ADN19868	ADN19868 Bacterial
31	52	57.8	712	8	ADS44332	ADS44332 Bacterial
32	52	57.8	828	4	AAW39674	AAW39674 Human pol
33	52	57.8	828	4	AAW39297	AAW39297 Human pro
34	52	57.8	828	6	ABR58327	ABR58327 NM 015046
35	52	57.8	828	8	ADP23318	ADP23318 PRO polyp
36	52	57.8	829	4	AAW80279	AAW80279 Human pro
37	52	57.8	829	4	AAW41460	AAW41460 Human pol
38	52	57.8	966	4	AAW79295	AAW79295 Human pro
39	52	57.8	2677	6	ABR58288	ABR58288 BCU0156 P
40	51	56.7	1417	4	ABR63329	ABR63329 Drosophi1
41	48	53.3	830	8	ADN21359	ADN21359 Bacterial
42	47.5	52.8	414	3	AAV77816	AAV77816 Yeast Dip
43	47.5	52.8	414	7	ABW01204	ABW01204 Saccharom
44	47.5	52.8	414	8	ADP44130	ADP44130 Yeast hel
45	47.5	52.8	683	8	ADS43834	ADS43834 Bacterial

## ALIGNMENTS

RESULT 1	ABW01182	standard; peptide: 21 AA.
ID	ABW01182	
XX	ABW01182;	
AC		
DT	15-JAN-2004 (first entry)	
XX		
DE	Saccharomyces cerevisiae motif VI peptide.	
XX		
KW	Modulator of translation termination; MTT1; helicase B; antiviral;	
XX	therapy; HCSB; nonsense mutation; yeast.	
OS	Saccharomyces cerevisiae.	
XX		
FT	Key	Location/Qualifiers
FT	Misc-difference 2	/label= Unknown
FT	Misc-difference 3	/note= "Xaa may be any amino acid"
FT	Misc-difference 10	/label= Unknown
FT	Misc-difference 16	/note= "Xaa may be any amino acid"
FT	Misc-difference 18	/label= Unknown
FT	Misc-difference 18	/note= "Xaa may be any amino acid"
FT	Misc-difference 20	/label= Unknown
FT	Misc-difference 20	/note= "Xaa may be any amino acid"
XX		
XX	US6630294-B1.	
XX	07-OCT-2003.	
XX	22-JUL-1999;	99US-00359268.
XX	22-JUL-1998;	98US-0093685P.
XX	(UNNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.	
XX	Peltz S, Czaplinski K, Dinman JD;	
XX	WPI; 2003-810549/76.	



XX Identifying an agent that increases nonsense suppression, for antiviral  
PT therapy, by contacting modulator of translation termination (Mtt1) in  
PT Saccharomyces cerevisiae with a test agent, and detecting specific  
PT binding to Mtt1.  
XX  
XX Disclosure; Col 43-44; 0pp; English.  
XX  
CC The invention relates to a method of identifying an agent that increases  
CC nonsense suppression, by contacting modulator of translation termination  
CC (Mtt1) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.  
CC The method is useful for identifying compositions or agents which  
CC increase nonsense suppression. The invention may also be used for  
CC antiviral therapy and for suppression of pathological nonsense mutations.  
CC The present sequence is Saccharomyces cerevisiae motif VI  
XX  
SQ Sequence 21 AA;  
Query Match 86.7%; Score 78; DB 7; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.7e-07;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LXXQYRMHPXISEFPYXXGL 21  
Db 1 LXXQYRMHPXISEFPYXXGL 21  
RESULT 2  
ADP44108  
ID ADP44108 standard; peptide; 21 AA.  
XX  
AC ADP44108;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Yeast translation termination modulation protein motif VI.  
XX  
XX gene therapy; translation termination; RNA helicase; MTT1;  
KM frameshift frequency; aberrant transcript degradation;  
KM peptidyl transferase modulation; beta-thalassemia; beta-globin;  
KM Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
KM Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;  
KM Ovarian Cancer; Wilms Tumour; Hirschprung disease; Cystic fibrosis;  
KM Kidney Stone; Familial hypercholesterolemia; Retinitis pigmentosa;  
KM Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; yeast.  
XX  
XX Saccharomyces cerevisiae.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FH MISC-difference 2..3  
FT MISC-difference 10 /note= "Any amino acid"  
FT MISC-difference 10 /note= "Any amino acid"  
FT MISC-difference 16 /note= "Any amino acid"  
FT MISC-difference 16 /note= "Any amino acid"  
FT MISC-difference 18 /note= "Any amino acid"  
FT MISC-difference 20 /note= "Any amino acid"  
FT MISC-difference 20 /note= "Any amino acid"  
XX  
XX US2004115787-A1.  
XX  
XX 17-JUN-2004.  
XX  
XX 28-AUG-2003; 2003US-00652334.  
XX  
XX 22-JUL-1998; 98US-0093685P.  
XX  
XX 22-JUL-1999; 99US-00359268.  
XX  
XX (PELTZ) PELTZ S.  
XX (CZAP/) CZAPLINSKI K.  
XX (DINM/) DINMAN J D.  
PA

XX  
XX Peltz S, Czaplinski K, Dinman JD;  
XX  
XX WPI; 2004-449400/42.  
DR  
XX  
XX Identifying a test composition or agent that modulates the efficiency of  
PT translation termination comprises contacting the MTT1 with the test  
PT composition or agent, and determining if the test composition or agent  
PT inhibits the MTT1.  
XX  
XX Claim 38; SEQ ID NO 6; 41pp; English.  
PS  
XX  
XX The invention relates to a method of identifying a test composition that  
CC modulates the efficiency of translation termination comprising contacting  
CC the RNA helicase MTT1 with a composition or agent under conditions  
CC permitting binding between the MTT1 and the composition, detecting  
CC specific binding of the test composition or agent to the MTT1, and  
CC determining if the test composition or agent inhibits the MTT1. The  
CC composition and methods are useful for modulating the fidelity of  
CC translation termination and for identifying agents that affect the  
CC functional activity of mRNAs by altering frameshift frequency, permit  
CC monitoring of a termination event, promote degradation of aberrant  
CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl  
CC transferase activity during initiation, elongation, termination and mRNA  
CC degradation of translation. The agents, which may be antagonists or  
CC agonists, are useful in screening, diagnostic and therapeutic purposes,  
CC for diseases or conditions resulting from or cause premature translation,  
CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular  
CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,  
CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,  
CC Hirschprung disease, Cystic fibrosis, Kidney Stones, Familial  
CC hypercholesterolemia, Retinitis pigmentosa, or Neurofibromatosis,  
CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents  
CC the amino acid sequence of the yeast translation termination modulation  
CC protein motif VI.  
XX  
SQ Sequence 21 AA;  
Query Match 86.7%; Score 78; DB 8; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.7e-07;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LXXQYRMHPXISEFPYXXGL 21  
Db 1 LXXQYRMHPXISEFPYXXGL 21  
RESULT 3  
AA77809  
ID AA77809 standard; peptide; 22 AA.  
XX  
AC AA77809;  
XX  
DT 31-MAY-2000 (first entry)  
XX  
DE Motif VI comprised in a gene modulating translation termination.  
XX  
XX Helicase B; HCSB; MTT1; modulator of translation termination; eRF1; eRF3;  
KM eukaryotic release factor; peptidyl transferase; beta-thalassemia;  
KM beta-globin; Duchenne/Becker Muscular Dystrophy; anti-leukemic.  
XX  
XX Unidentified.  
OS  
XX  
XX Key Location/Qualifiers  
FH MISC-difference 1..22 /note= "residues indicated Xaa are unspecified"  
FT MISC-difference 10 /note= "residues indicated Xaa are unspecified"  
FT MISC-difference 10 /note= "residues indicated Xaa are unspecified"  
XX  
XX WO200005586-A2.  
XX  
XX 03-FEB-2000.  
XX  
XX 22-JUL-1999; 99WO-US016602.  
XX



PR 22-JUL-1998; 98US-00120435.  
 XX (UYNE-) UNIV NEW JERSEY.  
 XX  
 XX Peltz S, Czaplinski K, Dimman JD;  
 XX WPI; 2000-171458/15.  
 DR  
 XX  
 PT New multiprotein complex which can modulate peptidyl transferase activity  
 PT during translation, useful to treat diseases associated with peptidyl  
 PT transferase activity e.g. Duchene/Becker Muscular Dystrophy.  
 PS  
 PS Claim 38; Page 79; 89pp; English.  
 XX  
 XX The invention provides a new multiprotein complex which can modulate  
 CC peptidyl transferase activity during translation. The complex comprises  
 CC the gene encoding Helicase B (HCSB, renamed MTT1, for Modulator of  
 CC Translation Termination) and the conserved proteins known to interact and  
 CC carry out translation termination in eukaryotic cells, peptidyl  
 CC eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to  
 CC modulate peptidyl transferase activity during translation in a cell. It  
 CC can be administered therapeutically combined with a carrier in  
 CC pharmaceutical compositions to treat diseases associated with peptidyl  
 CC transferase activity, especially diseases resulting from a nonsense or  
 CC frameshift mutation e.g. beta-thalassemia, beta-globin, Duchene/Becker  
 CC Muscular Dystrophy etc. It can be used to identify disease conditions  
 CC involving a defect in the complex, by transfecting cells with encoding  
 CC nucleic acid and determining the proportion of defective complex before  
 CC and after transfection. It is also useful to screen for drugs involved in  
 CC peptidyl transferase activity during translation, inhibiting the  
 CC interaction between MTT1 and eRF3 or involved in enhancing translation  
 CC termination. Vectors comprising polynucleotides encoding the complex (or  
 CC antisense sequences) can be constructed and introduced into cells to  
 CC interfere with complex expression and so modulate the efficiency of  
 CC translation termination of mRNA and/or degradation of aberrant  
 CC transcripts in a cell. Agents binding to the complex can be identified  
 CC and included in therapeutic compositions useful as above, and/or used to  
 CC modulate peptidyl transferase activity during translation in cells. They  
 CC are also useful to modulate the efficiency of translation termination of  
 CC mRNA at a nonsense codon and/or promote degradation of aberrant  
 CC transcripts in cells. The method can be used to identify agents/  
 CC compositions modulating binding to MTT1, useful to identify genes.  
 CC Sequences AAY77804-812 represent motifs 1-IX comprised in the genes of  
 CC interest, used for modulating translation termination  
 CC  
 XX  
 SQ Sequence 22 AA;  
 Query Match 86.7%; Score 78; DB 3; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-07;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LXXQYRMHPXISFPYXXGL 21  
 DB 1 LXXQYRMHPXISFPYXXGL 21  
 RESULT 4  
 ID AAY77813 standard; peptide; 415 AA.  
 XX  
 AC AAY77813;  
 XX  
 XX 31-MAY-2000 (first entry)  
 XX  
 DE Yeast Mtcl protein fragment.  
 XX  
 KM Helicase B; HCSB; MTT1; modulator of translation termination; eRF1; eRF3;  
 KM eukaryotic release factor; peptidyl transferase; beta-thalassemia;  
 KM beta-globin; Duchene/Becker Muscular Dystrophy; antenatal; yeast;  
 KM helicase.  
 XX  
 XX Saccharomyces cerevisiae.  
 XX

PN W0200005586-A2.  
 XX  
 PD 03-FEB-2000.  
 XX  
 XX 22-JUL-1999; 99WO-US016802.  
 PR  
 XX 22-JUL-1998; 98US-00120435.  
 FR  
 XX  
 XX (UYNE-) UNIV NEW JERSEY.  
 PA  
 PI Peltz S, Czaplinski K, Dimman JD;  
 XX WPI; 2000-171458/15.  
 DR  
 XX  
 PT New multiprotein complex which can modulate peptidyl transferase activity  
 PT during translation, useful to treat diseases associated with peptidyl  
 PT transferase activity e.g. Duchene/Becker Muscular Dystrophy.  
 PS  
 PS Example 1; Fig 1; 89pp; English.  
 XX  
 XX The invention provides a new multiprotein complex which can modulate  
 CC peptidyl transferase activity during translation. The complex comprises  
 CC the gene encoding Helicase B (HCSB, renamed MTT1, for Modulator of  
 CC Translation Termination) and the conserved proteins known to interact and  
 CC carry out translation termination in eukaryotic cells, peptidyl  
 CC eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to  
 CC modulate peptidyl transferase activity during translation in a cell. It  
 CC can be administered therapeutically combined with a carrier in  
 CC pharmaceutical compositions to treat diseases associated with peptidyl  
 CC transferase activity, especially diseases resulting from a nonsense or  
 CC frameshift mutation e.g. beta-thalassemia, beta-globin, Duchene/Becker  
 CC Muscular Dystrophy etc. It can be used to identify disease conditions  
 CC involving a defect in the complex, by transfecting cells with encoding  
 CC nucleic acid and determining the proportion of defective complex before  
 CC and after transfection. It is also useful to screen for drugs involved in  
 CC peptidyl transferase activity during translation, inhibiting the  
 CC interaction between MTT1 and eRF3 or involved in enhancing translation  
 CC termination. Vectors comprising polynucleotides encoding the complex (or  
 CC antisense sequences) can be constructed and introduced into cells to  
 CC interfere with complex expression and so modulate the efficiency of  
 CC translation termination of mRNA and/or degradation of aberrant  
 CC transcripts in a cell. Agents binding to the complex can be identified  
 CC and included in therapeutic compositions useful as above, and/or used to  
 CC modulate peptidyl transferase activity during translation in cells. They  
 CC are also useful to modulate the efficiency of translation termination of  
 CC mRNA at a nonsense codon and/or promote degradation of aberrant  
 CC transcripts in cells. The method can be used to identify agents/  
 CC compositions modulating binding to MTT1, useful to identify genes.  
 CC Sequences AAY77813-817 represent protein fragments from yeast superfamily  
 CC group I helicases  
 CC  
 XX  
 SQ Sequence 415 AA;  
 Query Match 75.0%; Score 67.5; DB 3; Length 415;  
 Best Local Similarity 62.5%; Pred. No. 0.00054;  
 Matches 15; Conservative 0; Mismatches 6; Indels 3; Gaps 1;  
 QY 1 LXXQYRMHPXISFP---YXXGL 21  
 DB 213 LDTQYRMHPXISFPKIKYNGEL 236  
 RESULT 5  
 ID ABW01201 standard; protein; 415 AA.  
 XX  
 AC ABW01201;  
 XX  
 XX 15-JAN-2004 (first entry)  
 DT  
 XX Saccharomyces cerevisiae modulator of translation termination protein.  
 DE  
 XX Modulator of translation termination; MTT1; helicase B; antiviral;  
 KM

KM therapy; HCSB; nonsense mutation; yeast.  
 XX  
 OS Saccharomyces cerevisiae.  
 XX  
 PN US6630294-B1.  
 XX  
 PD 07-OCT-2003.  
 XX  
 PR 22-JUL-1999; 99US-00359268.  
 XX  
 PR 22-JUL-1998; 98US-0093685P.  
 XX  
 PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.  
 XX  
 PI Peltz S, Czaplinski K, Dimman JD;  
 XX  
 DR WPI; 2003-810549/76.  
 XX  
 PT Identifying an agent that increases nonsense suppression, for antiviral  
 PT therapy, by contacting modulator of translation termination (Mtt1) in  
 PT Saccharomyces cerevisiae with a test agent, and detecting specific  
 PT binding to Mtt1.  
 XX  
 PS Disclosure; Col 49-52; 0pp; English.  
 XX  
 CC The invention relates to a method of identifying an agent that increases  
 CC nonsense suppression, by contacting modulator of translation termination  
 CC (Mtt1) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.  
 CC The method is useful for identifying compositions or agents which  
 CC increase nonsense suppression. The invention may also be used for  
 CC antiviral therapy and for suppression of pathological nonsense mutations.  
 CC The present sequence is Saccharomyces cerevisiae MTT1 protein  
 CC  
 XX Sequence 415 AA;  
 SQ  
 Query Match 75.0%; Score 67.5; DB 7; Length 415;  
 Best Local Similarity 62.5%; Pred. No. 0.00054;  
 Matches 15; Conservative 0; Mismatches 6; Indels 3; Gaps 1;  
 QY 1 LXXQYRMHPKISSEFP--XXYXGL 21  
 Db 213 LDTQYRMHPKISSEFPKIKYNGEL 236

RESULT 6  
 ADP44127  
 ID ADP44127 standard; protein; 415 AA.  
 XX  
 AC ADP44127;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Yeast helicase Mtt1.  
 XX  
 KW gene therapy; translation termination; RNA helicase; MTT1;  
 KW frameshift frequency; aberrant transcript degradation;  
 KW peptidyl transferase modulation; beta-thalassemia; beta-globin;  
 KW Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
 KW Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;  
 KW Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic fibrosis;  
 KW Kidney Stone; Familial hypercholesterolaemia; Retinitis pigmentosa;  
 KW Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; yeast; enzyme.  
 XX  
 OS Saccharomyces cerevisiae.  
 XX  
 PN US2004115787-A1.  
 XX  
 PD 17-JUN-2004.  
 XX  
 PR 28-AUG-2003; 2003US-00652334.  
 XX  
 PR 22-JUL-1998; 98US-0093685P.  
 XX  
 PR 22-JUL-1999; 99US-00359268.

XX  
 PA (PELTZ) PELTZ S.  
 PA (CZAP) CZAPLINSKI K.  
 PA (DINM) DINMAN J D.  
 XX  
 PI Peltz S, Czaplinski K, Dimman JD;  
 XX  
 DR WPI; 2004-449400/42.  
 XX  
 PR Identifying a test composition or agent that modulates the efficiency of  
 PT translation termination comprises contacting the MTT1 with the test  
 PT composition or agent, and determining if the test composition or agent  
 PT inhibits the MTT1.  
 XX  
 PS Disclosure; SEQ ID NO 25; 41pp; English.  
 XX  
 CC The invention relates to a method of identifying a test composition that  
 CC modulates the efficiency of translation termination comprising contacting  
 CC the RNA helicase MTT1 with a composition or agent under conditions  
 CC permitting binding between the MTT1 and the composition, detecting  
 CC specific binding of the test composition or agent to the MTT1, and  
 CC determining if the test composition or agent inhibits the MTT1. The  
 CC composition and methods are useful for modulating the fidelity of  
 CC translation termination or for identifying agents that affect the  
 CC functional activity of mRNAs by altering frameshift frequency, permit  
 CC monitoring of a termination event, promote degradation of aberrant  
 CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl  
 CC transferase activity during initiation, elongation, termination and mRNA  
 CC degradation of translation. The agents, which may be antagonists or  
 CC agonists, are useful in screening, diagnostic and therapeutic purposes,  
 CC for diseases or conditions resulting from or cause premature translation,  
 CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular  
 CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,  
 CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,  
 CC Hirschsprung disease, Cystic fibrosis, Kidney Stones, Familial  
 CC Hypercholesterolaemia, Retinitis pigmentosa, or Neurofibromatosis.  
 CC The present sequence represents the amino acid sequence of the yeast helicase Mtt1.  
 CC  
 XX Sequence 415 AA;  
 SQ  
 Query Match 75.0%; Score 67.5; DB 8; Length 415;  
 Best Local Similarity 62.5%; Pred. No. 0.00054;  
 Matches 15; Conservative 0; Mismatches 6; Indels 3; Gaps 1;  
 QY 1 LXXQYRMHPKISSEFP--XXYXGL 21  
 Db 213 LDTQYRMHPKISSEFPKIKYNGEL 236

RESULT 7  
 ADN19434  
 ID ADN19434 standard; protein; 925 AA.  
 XX  
 AC ADN19434;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE Bacterial polypeptide #2087.  
 XX  
 KW Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.  
 XX  
 OS Bacteria.  
 XX  
 PN US2003233675-A1.  
 XX  
 PD 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.  
PF 21-FEB-2002; 2002US-0360039P.  
XX  
XX (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX WPI; 2004-061375/06.  
XX  
XX New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
XX Claim 1; SEQ ID NO 2087; 122pp; English.  
XX  
XX The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition. Improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at seqdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 925 AA;  
XX  
XX Query Match 72.8%; Score 65.5; DB 8; Length 925;  
Best Local Similarity 58.3%; Pred. No. 0.0033;  
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;  
XX  
QY 1 LXXQYRMHPXISEFPX---YXGXL 21  
| | | | | : | | | | |  
Db 626 LVVQYRMHPCLSEFPNMFYEGTL 649  
XX  
XX RESULT 8  
AD523757 ID AD523757 standard; protein; 992 AA.  
XX  
XX AC AD523757;  
XX  
XX 02-DEC-2004 (first entry)  
XX  
XX Bacterial polypeptide #12790.  
XX  
XX Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

KM bacterial polypeptide.  
XX  
XX Bacteria.  
XX  
XX US2003233675-A1.  
XX  
XX 18-DEC-2003.  
XX  
XX 20-FEB-2003; 2003US-00369493.  
XX  
XX 21-FEB-2002; 2002US-0360039P.  
XX  
XX (CAOY/) CAO Y.  
XX (HINK/) HINKLE G J.  
XX (SLAT/) SLATER S C.  
XX (CHEN/) CHEN X.  
XX (GOLD/) GOLDMAN B S.  
XX  
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX WPI; 2004-061375/06.  
XX  
XX New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
XX Claim 1; SEQ ID NO 12790; 122pp; English.  
XX  
XX The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition. Improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at seqdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 992 AA;  
XX  
XX Query Match 72.8%; Score 65.5; DB 8; Length 992;  
Best Local Similarity 58.3%; Pred. No. 0.0036;  
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;  
XX  
QY 1 LXXQYRMHPXISEFPX---YXGXL 21  
| | | | | : | | | | |  
Db 566 LVVQYRMHPCLSEFPNMFYEGSL 569  
XX  
XX RESULT 9  
AAV77815 ID AAV77815 standard; peptide; 471 AA.  
XX  
XX AC AAV77815;  
XX  
XX 31-MAY-2000 (first entry)  
XX  
XX Yeast Sen1 protein fragment.



PN US2004115787-A1.  
 XX 17-JUN-2004.  
 PD 28-AUG-2003; 2003US-00652334.  
 XX 22-JUL-1998; 98US-0093685P.  
 PR 22-JUL-1999; 99US-00359268.  
 XX (PELT/) PELTZ S.  
 PA (CZAP/) CZAPLINSKI K.  
 PA (DINM/) DINMAN J D.  
 PI Peltz S, Czaplinski K, Dinman JD;  
 XX WPI; 2004-449400/42.  
 DR WPI; 2004-449400/42.  
 PT Identifying a test composition or agent that modulates the efficiency of  
 PT translation termination comprises contacting the MT1 with the test  
 PT composition or agent, and determining if the test composition or agent  
 PT inhibits the MT1.  
 PS Disclosure; SEQ ID NO 26; 41pp; English.  
 XX The invention relates to a method of identifying a test composition that  
 CC modulates the efficiency of translation termination comprising contacting  
 CC the RNA helicase MT1 with a composition or agent under conditions  
 CC permitting binding between the MT1 and the composition, detecting  
 CC specific binding of the test composition or agent to the MT1, and  
 CC determining if the test composition or agent inhibits the MT1. The  
 CC composition and methods are useful for modulating the fidelity of  
 CC translation termination or for identifying agents that affect the  
 CC functional activity of mRNAs by altering frameshift frequency, permit  
 CC monitoring of a termination event, promote degradation of aberrant  
 CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl  
 CC transferase activity during initiation, elongation, termination and mRNA  
 CC degradation of translation. The agents, which may be antagonists or  
 CC agonists, are useful in screening, diagnostic and therapeutic purposes,  
 CC for diseases or conditions resulting from or cause premature translation,  
 CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular  
 CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,  
 CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,  
 CC Hirschsprung disease, Cystic fibrosis, Kidney Stones, Familial  
 CC hypercholesterolemia, Retinitis Pigmentosa, or Neurofibromatosis,  
 CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents  
 CC the amino acid sequence of the yeast helicase Sen1.  
 CC  
 SQ Sequence 472 AA;  
 Query Match 71.7%; Score 64.5; DB 8; Length 472;  
 Best Local Similarity 58.3%; Pred. No. 0.0023; Indels 3; Gaps 1;  
 Matches 14; Conservative 1; Mismatches 6;  
 QY 1 LXXQYRMHPXISRPX---YXGXL 21  
 | | | | | | | | | | | | | | | | | | | | | |  
 DB 301 LDVQYRMHPXISKRPSPSEFYGRL 324  
 RESULT 12  
 ABR53451  
 ID ABR53451 standard; protein; 2000 AA.  
 XX ABR53451;  
 AC ABR53451;  
 XX 20-JUN-2003 (first entry)  
 DT 20-JUN-2003 (first entry)  
 XX Protein sequence #SEQ ID 1767.  
 DE Protein sequence #SEQ ID 1767.  
 XX  
 KM Multiprotein complex; eukaryote; drug target; diagnosis.  
 XX Saccharomycetaceae cerevisiae.  
 OS Saccharomycetaceae cerevisiae.  
 XX  
 PI BPI258494-A1.

XX 20-NOV-2002.  
 PD 20-DEC-2001; 2001EP-00130253.  
 XX 15-MAY-2001; 2001EP-00111774.  
 PR (CELL-) CELLZOMB AG.  
 PA (CELL-) CELLZOMB AG.  
 PI Bauer A, Gavin A, Grandt P, Krause R, Kruse UD, Kuester BD;  
 PI Marzloch M, Schultz JD, Superti-Furga GD;  
 DR WPI; 2003-250078/25.  
 DR N-PSDB; ACC61493.  
 XX  
 PT New isolated protein complexes useful for diagnosing a disease or  
 PT disorder, or as a target for an active agent of a pharmaceutical,  
 PT preferably a drug target in the treatment or prevention of disease or  
 PT disorder.  
 PS Disclosure; SEQ ID NO 1767; 17pp + Sequence Listing; English.  
 XX The invention relates to multiprotein complexes from eukaryotes. Proteins  
 CC of the invention and DNA sequences encoding them are given in records  
 CC ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are  
 CC obtainable by using a protein as a bait and isolating the set of proteins  
 CC which is attached thereto from cells. Such protein complexes may comprise  
 CC up to 30 distinct proteins. Protein complexes of the invention are useful  
 CC for diagnosing a disease or disorder, or as a target for an active agent  
 CC of a pharmaceutical, preferably a drug target in the treatment or  
 CC prevention of a disease or disorder. Note: The sequence data for this  
 CC patent is not represented in the printed specification, but is based on  
 CC sequence information supplied by the European Patent Office. The complete  
 CC document is available on CD-ROM  
 XX  
 SQ Sequence 2000 AA;  
 Query Match 71.7%; Score 64.5; DB 6; Length 2000;  
 Best Local Similarity 58.3%; Pred. No. 0.013; Indels 3; Gaps 1;  
 Matches 14; Conservative 1; Mismatches 6;  
 QY 1 LXXQYRMHPXISRPX---YXGXL 21  
 | | | | | | | | | | | | | | | | | | | | | |  
 DB 1651 LDVQYRMHPXISKRPSPSEFYGRL 1674  
 RESULT 13  
 ADK64622  
 ID ADK64622 standard; protein; 2000 AA.  
 XX ADK64622;  
 AC ADK64622;  
 XX 06-MAY-2004 (first entry)  
 DT 06-MAY-2004 (first entry)  
 XX Disease treating protein complex-derived protein #1064.  
 DE Disease treating protein complex-derived protein #1064.  
 XX protein complex; drug target; diagnosis.  
 KM Unidentified.  
 OS Unidentified.  
 XX EP1336608-A2.  
 PN EP1336608-A2.  
 XX 27-AUG-2003.  
 PD 20-DEC-2002; 2002EP-00102902.  
 PP 20-DEC-2002; 2002EP-00102902.  
 XX 20-DEC-2001; 2001EP-00130253.  
 PR (CELL-) CELLZOMB AG.  
 PA (CELL-) CELLZOMB AG.  
 PI Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;  
 PI Marzloch M, Grandt P, Krause R, Kruse U, Merino A, Bauch A;  
 PI Michon A, Leutwein C, Rick J;

XX WPI: 2003-638460/61.  
DR N-PSDB; ADK64623.  
XX  
PT New proteins and protein complexes from eukaryotes, useful as targets in  
PT drug screening, or in diagnosing or screening for the presence of a  
PT disease or disorder, or a predisposition for developing a disease or  
PT disorder in a subject.  
XX  
PS Disclosure; SEQ ID NO 2127; 13pp; English.  
XX  
CC The invention relates to novel protein complexes comprising a first and a  
CC second protein, or its derivative, fragment, homologue or variant. The  
CC proteins are selected from given protein complexes, which are not defined  
CC in the specification. The variants are encoded by nucleic acids that  
CC hybridize to the nucleic acids encoding the proteins under low stringency  
CC conditions. The protein complexes are useful as targets for an active  
CC agent of a pharmaceutical. These protein complexes are particularly  
CC useful as drug targets for the treatment or preventing of a disease or  
CC disorder. The complexes and methods above are useful in diagnosing or  
CC screening for the presence of a disease or disorder or a predisposition  
CC for developing a disease or disorder in a subject. These are also useful  
CC in screening for a drug for treatment or prevention of a disease or  
CC disorder. The molecule that modulates the amount, activity or protein  
CC components of the complex is useful for the manufacture of a medicament  
CC for the treatment or prevention of a disease or disorder. This sequence  
CC corresponds to a protein of the invention. (Note: the sequence data for  
CC this patent did not form part of the printed specification but was  
CC obtained from the EPO in electronic format).  
XX  
SQ Sequence 2000 AA;  
XX  
QY Query Match 71.7%; Score 64.5; DB 7; Length 2000;  
Best Local Similarity 58.3%; Pred. No. 0.013; Mismatches 6; Indels 3; Gaps 1;  
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;  
QY 1 LXXQYRMHPXISEFPX--YXGXL 21  
Db 1651 LDVQYRMHPXISKPFSSSEFYQRL 1674  
XX  
RESULT 14  
ADN19177  
ID ADN19177 standard; protein; 2231 AA.  
XX  
AC ADN19177;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Bacterial polypeptide #1830.  
XX  
KM Recombinant DNA construct; transformed plant; improved plant property;  
KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KM pathogen tolerance; pest tolerance; plant disease resistance;  
KM cell cycle pathway modification; plant growth regulator;  
KM homologous recombination; seed oil yield; protein yield; carbohydrate;  
KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KM bacterial polypeptide.  
XX  
OS Bacteria.  
XX  
PN US2003233675-A1.  
XX  
PD 18-DEC-2003.  
XX  
PF 20-FEB-2003; 2003US-00369493.  
XX  
PR 21-FEB-2002; 2002US-0360039P.  
XX  
CAO Y.  
PA (HINKLE) HINKLE G J.  
PA (SLATER) SLATER S C.  
PA (CHEN) CHEN X.

PA (GOLD/) GOLDMAN B S.  
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX  
DR WPI: 2004-061375/06.  
XX  
PT New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
PS Claim 1; SEQ ID NO 1830; 122pp; English.  
XX  
CC The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX  
SQ Sequence 2231 AA;  
XX  
QY Query Match 71.7%; Score 64.5; DB 8; Length 2231;  
Best Local Similarity 58.3%; Pred. No. 0.014; Mismatches 6; Indels 3; Gaps 1;  
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;  
QY 1 LXXQYRMHPXISEFPX--YXGXL 21  
Db 1651 LDVQYRMHPXISKPFSSSEFYQRL 1674  
XX  
RESULT 15  
ABO58160  
ID ABO58160 standard; protein; 71 AA.  
XX  
AC ABO58160;  
XX  
DT 29-JUN-2004 (first entry).  
XX  
DE Human genome derived single exon protein #4394.  
XX  
KM Human; gene expression; single exon probe; microarray;  
KM alternative splicing event; genomic alteration.  
XX  
OS Homo sapiens.  
XX  
PN US2003194704-A1.  
XX  
PD 16-OCT-2003.  
XX  
PF 03-APR-2002; 2002US-00029386.  
XX  
PR 03-APR-2002; 2002US-00029386.  
XX  
CAO Y.  
PA (PENN/) PENN S G.  
PA (RANK/) RANK D R.

PA (HANZ/) HANZEL D K.

PI Penn SG, Rank DR, Hanzel DK;

XX WPI; 2004-119264/12.

DR  
XX  
PT New human genome-derived single exon nucleic acid probes useful for human  
PT gene expression analysis, for identifying or characterizing alternative  
PT splicing events, for assessing genomic alterations or as tools for  
PT surveying tissues.

XX  
PS Claim 45; SEQ ID NO 31794; 80pp; English.

XX  
CC The invention relates to a nucleic acid probe for measuring human gene  
CC expression, comprising any of the 27,400 fully defined nucleotide  
CC sequences in the specification, or their complements or fragments, and  
CC encoding at least 8 amino acids of any of the 688 amino acid sequences  
CC fully defined in the specification. The probe is a single exon probe that  
CC hybridizes under high stringency conditions to a nucleic acid molecule  
CC expressed in human cells or tissues. Also included are a spatially-  
CC addressable set of single exon nucleic acid probes for measuring human  
CC gene expression (comprising a plurality of single exon nucleic acid  
CC probes cited above, where each of the plurality of probes is separately  
CC and addressably isolatable or amplifiable from the plurality), a single  
CC exon microarray for measuring human gene expression, a method of  
CC measuring human gene expression, a vector comprising the single exon  
CC probe cited above, an ORF-encoded peptide comprising at least 8  
CC contiguous amino acids of any of the above-mentioned amino acid  
CC sequences (optionally with conservative amino acid substitutions), an  
CC isolated antibody that binds specifically to a peptide cited above,  
CC methods of selling and/or licensing single exon probes or microarrays to  
CC a customer desiring to measure gene expression, a method of providing  
CC human gene expression data by subscription, and a computer-readable  
CC storage medium which contains a database having a plurality of records  
CC (each record including data on the expression of a single exon probe  
CC cited above. The probe, methods and apparatus are useful in gene  
CC expression analysis. The probes may be used as tools for surveying  
CC tissues to detect the presence of expressed messages that contain their  
CC specific exon, or in constructing genome-derived single exon microarrays.  
CC In addition, the probes are used in identifying and characterizing  
CC alternative splicing events, in detecting and characterizing gross  
CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe protein of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030194704

XX  
SQ Sequence 71 AA;

Query Match 68.3%; Score 61.5; DB 8; Length 71;

Best Local Similarity 54.2%; Pred. No. 0.00093;  
Matches 13; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 1 LXXOYRMHPXISRPX---YXGXL 21  
| | | | | : | | | | |  
DB 42 LQVQYRMHPALSAFPSPNIFREGSL 65

Search completed: April 18, 2005, 08:03:48  
Job time : 104.651 secs

This page blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 07:37:21 / Search time 26.2117 Seconds  
(without alignments)  
59.807 Million cell updates/sec

Title: US-10-652-334-6  
Perfect score: 90  
Sequence: 1 LXXQYRMHPXISBPYXGXL 21

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pcp:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pcp:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pcp:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pcp:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pcp:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pcp:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	86.7	21	US-09-359-268A-6	Sequence 6, Appl1
2	67.5	75.0	415	US-09-359-268A-25	Sequence 25, Appl1
3	67.5	75.0	917	US-09-248-796A-19347	Sequence 19347, A
4	65.5	72.8	106	US-09-248-796A-18134	Sequence 18134, A
5	64.5	71.7	472	US-09-359-268A-26	Sequence 26, Appl1
6	61.5	68.3	1043	US-08-724-354D-4	Sequence 4, Appl1
7	61.5	68.3	1043	US-08-724-354D-4	Sequence 4, Appl1
8	61.5	68.3	1118	US-08-724-354D-2	Sequence 2, Appl1
9	61.5	68.3	1140	US-09-949-016-10116	Sequence 10116, A
10	61.5	68.3	1140	US-09-949-016-10116	Sequence 29, Appl1
11	59.5	66.1	380	US-09-359-268A-29	Sequence 22, Appl1
12	59.5	66.1	971	US-08-724-354D-22	Sequence 22, Appl1
13	59.5	66.1	971	US-09-270-984A-22	Sequence 28, Appl1
14	59.5	66.1	971	US-09-177-431-8	Sequence 28, Appl1
15	47.5	52.8	414	US-09-359-268A-28	Sequence 483, App
16	47.5	52.8	683	US-09-538-092-483	Sequence 16148, A
17	44.5	49.4	211	US-09-902-540-16148	Sequence 58044, A
18	44.5	49.4	102	US-09-270-767-58044	Sequence 42726, A
19	44.5	49.4	152	US-09-270-767-42726	Sequence 17548, A
20	43	47.8	1037	US-09-252-991A-17548	Sequence 18131, A
21	42	46.7	426	US-09-248-796A-15170	Sequence 11607, A
22	41.5	46.1	413	US-09-248-796A-15170	Sequence 20958, A
23	41.5	46.1	821	US-09-902-540-11607	Sequence 57996, A
24	41.5	46.1	307	US-09-248-796A-20958	Sequence 42678, A
25	40	44.4	47	US-09-270-767-57996	Sequence 5447867-4
26	40	44.4	235	US-09-270-767-42678	Patent No. 5447867
27	40	44.4	389	5447867-4	

28	40	44.4	389	6	5447867-4	Patent No. 5447867
29	40	44.4	773	1	US-08-019-870-1	Sequence 1, Appl1
30	40	44.4	773	1	US-08-019-870-6	Sequence 6, Appl1
31	40	44.4	774	1	US-07-747-901A-3	Sequence 3, Appl1
32	40	44.4	774	1	US-07-935-312-3	Sequence 3, Appl1
33	40	44.4	774	1	US-08-019-870-3	Sequence 3, Appl1
34	40	44.4	774	1	US-08-019-870-5	Sequence 5, Appl1
35	40	44.4	774	1	US-08-019-870-5	Sequence 8, Appl1
36	40	44.4	774	1	US-08-019-870-11	Sequence 11, Appl1
37	40	44.4	774	1	US-08-314-309A-21	Sequence 21, Appl1
38	40	44.4	774	1	US-08-633-760-44	Sequence 44, Appl1
39	40	44.4	774	1	US-08-633-760-46	Sequence 46, Appl1
40	40	44.4	774	1	US-08-633-760-48	Sequence 48, Appl1
41	40	44.4	774	1	US-08-633-760-50	Sequence 50, Appl1
42	40	44.4	774	1	US-08-633-760-52	Sequence 52, Appl1
43	39	43.3	62	4	US-09-621-976-6904	Sequence 6904, Ap
44	39	43.3	306	4	US-09-248-796A-16127	Sequence 16127, A
45	38	42.2	343	4	US-09-248-796A-19126	Sequence 19126, A

## ALIGNMENTS

```
RESULT 1
US-09-359-268A-6
Sequence 6, Application US/09359268A
Patent No. 6630294
GENERAL INFORMATION:
APPLICANT: Peltz, Stuart
APPLICANT: Czaplinski, Kevin
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 601-1-85N
CURRENT APPLICATION NUMBER: US/09/359,268A
CURRENT FILING DATE: 1999-07-22
PRIOR APPLICATION NUMBER: 60/093,685
PRIOR FILING DATE: 1998-07-22
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 6
LENGTH: 21
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
FEATURES:
OTHER INFORMATION: Xaa = any amino acid
US-09-359-268A-6
Query Match 86.7%; Score 78; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 LXXQYRMHPXISBPYXGXL 21
Db 1 LXXQYRMHPXISBPYXGXL 21
RESULT 2
US-09-359-268A-25
Sequence 25, Application US/09359268A
Patent No. 6630294
GENERAL INFORMATION:
APPLICANT: Peltz, Stuart
APPLICANT: Czaplinski, Kevin
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 601-1-85N
CURRENT APPLICATION NUMBER: US/09/359,268A
CURRENT FILING DATE: 1999-07-22
PRIOR APPLICATION NUMBER: 60/093,685
```

PRIOR FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 25  
LENGTH: 415  
TYPE: PRT  
ORGANISM: saccharomyces cerevisiae  
US-09-359-268A-25

Query Match 75.0%; Score 67.5; DB 4; Length 415;  
Best Local Similarity 62.5%; Pred. No. 0.00025;  
Matches 15; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

OY 1 LXXOYRMHPXISEPFX---YXXGL 21  
DB 213 LDYOYRMHPXISEPFXIKYNGEL 236

RESULT 3  
US-09-248-796A-19347  
Sequence 19347, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Kelch Weinstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 19347  
LENGTH: 917  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-19347

Query Match 75.0%; Score 67.5; DB 4; Length 917;  
Best Local Similarity 62.5%; Pred. No. 0.00059;  
Matches 15; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

OY 1 LXXOYRMHPXISEPFX---YXXGL 21  
DB 682 LDYOYRMHPXISEPFXRYGGL 705

RESULT 4  
US-09-248-796A-18134  
Sequence 18134, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Kelch Weinstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 18134  
LENGTH: 106  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-18134

Query Match 72.8%; Score 65.5; DB 4; Length 106;  
Best Local Similarity 58.3%; Pred. No. 0.00013;

Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;  
OY 1 LXXOYRMHPXISEPFX---YXXGL 21  
DB 24 LDYOYRMHPXISEPFXNMFYEGSL 47

RESULT 5  
US-09-359-268A-26  
Sequence 26, Application US/09359268A  
Patent No. 6630394  
GENERAL INFORMATION:  
APPLICANT: Peltz, Stuart  
APPLICANT: Czapinski, Kevin  
APPLICANT: Diman, Jonathan D.  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USRS  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/09/359,268A  
CURRENT FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
PRIOR FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 26  
LENGTH: 472  
TYPE: PRT  
ORGANISM: saccharomyces cerevisiae  
US-09-359-268A-26

Query Match 71.7%; Score 64.5; DB 4; Length 472;  
Best Local Similarity 58.3%; Pred. No. 0.001;  
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

OY 1 LXXOYRMHPXISEPFX---YXXGL 21  
DB 301 LDYOYRMHPXISEPFXSYOGRL 324

RESULT 6  
US-08-724-354D-4  
Sequence 4, Application US/08724354D  
Patent No. 5994119  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,354D  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/016,482  
FILING DATE: 29-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hallie, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070

TELEFAX: 619-678-5099  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1043 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-724-354D-4

Query Match 68.3%; Score 61.5; DB 2; Length 1043;  
Best Local Similarity 54.2%; Pred. No. 0.0084;  
Matches 13; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 1 LXXOYRMPXISEPPX---YXGXL 21  
Db 621 LQVQYRMPALSAFSPNIFYEGSL 644

RESULT 7  
US-09-270-984A-4  
; Sequence 4, Application US/09270984A  
; Patent No. 6048965

GENERAL INFORMATION:  
; APPLICANT: Dietz, Harry C.  
; TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
; TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/270.984A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/724.354  
; FILING DATE:

ATTORNEY/AGENT INFORMATION:  
; NAME: Halle, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/090001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-678-5070  
; TELEFAX: 619-678-5099

INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1043 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-270-984A-4

Query Match 68.3%; Score 61.5; DB 3; Length 1043;  
Best Local Similarity 54.2%; Pred. No. 0.0084;  
Matches 13; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 1 LXXOYRMPXISEPPX---YXGXL 21  
Db 621 LQVQYRMPALSAFSPNIFYEGSL 644

RESULT 8  
US-08-724-354D-2  
; Sequence 2, Application US/08724354D  
; Patent No. 5994119

GENERAL INFORMATION:  
; APPLICANT: Dietz, Harry C.  
; TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
; TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724.354D  
; FILING DATE: 01-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/016,482  
; FILING DATE: 29-APR-1996

ATTORNEY/AGENT INFORMATION:  
; NAME: Halle, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/090001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-678-5070  
; TELEFAX: 619-678-5099

INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1118 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-724-354D-2

Query Match 68.3%; Score 61.5; DB 2; Length 1118;  
Best Local Similarity 54.2%; Pred. No. 0.0091;  
Matches 13; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 1 LXXOYRMPXISEPPX---YXGXL 21  
Db 698 LQVQYRMPALSAFSPNIFYEGSL 721

RESULT 9  
US-09-270-984A-2  
; Sequence 2, Application US/09270984A  
; Patent No. 6048965

GENERAL INFORMATION:  
; APPLICANT: Dietz, Harry C.  
; TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
; TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/270.984A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/724.354

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1118 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-270-984A-2

Query Match 68.3%; Score 61.5; DB 3; Length 1118;  
Best Local Similarity 54.2%; Pred. No. 0.0091;  
Matches 13; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

OY 1 LXXQYRMHPXISEFPX---YXGXL 21  
DB 698 LQVQYRMHPALSAFSPNIFYEGSL 721

RESULT 10  
US-09-949-016-10116  
Sequence 10116, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CLO01307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10116  
LENGTH: 1140  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-10116

Query Match 68.3%; Score 61.5; DB 4; Length 1140;  
Best Local Similarity 54.2%; Pred. No. 0.0093;  
Matches 13; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

OY 1 LXXQYRMHPXISEFPX---YXGXL 21  
DB 720 LQVQYRMHPALSAFSPNIFYEGSL 743

RESULT 11  
US-09-359-268A-29  
Sequence 29, Application US/09359268A  
Patent No. 6630294  
GENERAL INFORMATION:  
APPLICANT: Peltz, Stuart  
APPLICANT: Czaplinski, Kevin  
APPLICANT: Dimman, Jonathan D.  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/09/359,268A  
CURRENT FILING DATE: 1999-07-22

PRIOR APPLICATION NUMBER: 60/093,685  
PRIOR FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 29  
LENGTH: 380  
TYPE: PRT  
ORGANISM: saccharomyces cerevisiae  
US-09-359-268A-29

Query Match 66.1%; Score 59.5; DB 4; Length 380;  
Best Local Similarity 54.2%; Pred. No. 0.0065;  
Matches 13; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

OY 1 LXXQYRMHPXISEFPX---YXGXL 21  
DB 211 LEVQYRMHPYLSFSPNMFYEGSL 234

RESULT 12  
US-08-724-354D-22  
Sequence 22, Application US/08724354D  
Patent No. 5994119  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,354D  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/016,482  
FILING DATE: 29-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 971 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-724-354D-22

Query Match 66.1%; Score 59.5; DB 2; Length 971;  
Best Local Similarity 54.2%; Pred. No. 0.018;  
Matches 13; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

OY 1 LXXQYRMHPXISEFPX---YXGXL 21  
DB 634 LEVQYRMHPYLSFSPNMFYEGSL 657

RESULT 13  
US-09-270-984A-22  
Sequence 22, Application US/09270984A

Patent No. 6048965  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/270,984A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/724,354  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Haille, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 971 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-270-984A-22

Query Match 66.1%; Score 59.5; DB 3; Length 971;  
Best Local Similarity 54.2%; Pred. No. 0.018;  
Matches 13; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 1 LXXQYRMHPXISEFPX---YXGXL 21  
DB 634 LEVQYRMNPLYSEFPSPNMFYEGSL 657

RESULT 14  
US-09-177-431-8  
Sequence 8, Application US/09177431  
Patent No. 6071700  
GENERAL INFORMATION:  
APPLICANT: He, Feng  
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE  
TITLE OF INVENTION: ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/177,431  
FILING DATE:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/955,472  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pasase, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 07917/050001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-9806  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 971 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
US-09-177-431-8

Query Match 66.1%; Score 59.5; DB 3; Length 971;  
Best Local Similarity 54.2%; Pred. No. 0.018;  
Matches 13; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 1 LXXQYRMHPXISEFPX---YXGXL 21  
DB 634 LEVQYRMNPLYSEFPSPNMFYEGSL 657

RESULT 15  
US-09-359-268A-28  
Sequence 28, Application US/09359268A  
Patent No. 6630294  
GENERAL INFORMATION:  
APPLICANT: Peltz, Stuart  
APPLICANT: Caplinet, Kevin  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/09/359,268A  
CURRENT FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
PRIOR FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 28  
LENGTH: 414  
TYPE: PRT  
ORGANISM: saccharomyces cerevisiae  
US-09-359-268A-28

Query Match 52.8%; Score 47.5; DB 4; Length 414;  
Best Local Similarity 50.0%; Pred. No. 1.1;  
Matches 12; Conservative 1; Mismatches 8; Indels 3; Gaps 1;

QY 1 LXXQYRMHPXISEFPX---YXGXL 21  
DB 238 LNVQYRMNOKIMFPPSHSMYNGRL 261

Search completed: April 18, 2005, 08:18:35  
Job time: 27.2117 secs

This Page Blank (uspb)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 08:06:16 ; Search time 74.4964 Seconds  
(without alignments)  
93.693 Million cell updates/sec

Title: US-10-652-334-6  
Sequence: 1 LXXQYRHPXISEFPYXGXL 21

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications RA:\*

- 1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubppaa/US09C\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubppaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	86.7	21	US-10-652-334-6	Sequence 6, Appl1
2	67.5	75.0	415	US-10-652-334-25	Sequence 25, Appl1
3	65.5	72.8	199	US-10-767-701-38822	Sequence 38822, A
4	65.5	72.8	734	US-10-424-599-205843	Sequence 205843, A
5	65.5	72.8	925	US-10-369-493-2087	Sequence 2087, Ap
6	64.5	71.7	925	US-10-369-493-12790	Sequence 12790, A
7	64.5	71.7	472	US-10-652-334-26	Sequence 26, Appl
8	64.5	71.7	2231	US-10-369-493-1830	Sequence 1830, Ap
9	63.5	70.6	163	US-10-767-701-49859	Sequence 49859, A
10	62.5	69.4	638	US-10-437-963-181890	Sequence 181890, A
11	61.5	68.3	71	US-10-029-386-31794	Sequence 31794, A
12	61.5	68.3	935	US-10-408-765A-1706	Sequence 1706, Ap
13	61.5	68.3	1118	US-10-474-553-6	Sequence 6, Appl1

14	60.5	67.2	219	US-10-767-701-57601	Sequence 57601, A
15	60	66.7	1975	US-10-437-963-140079	Sequence 14079, A
16	59.5	66.1	380	US-10-652-334-29	Sequence 29, Appl
17	59.5	66.1	559	US-10-437-963-190072	Sequence 190072, A
18	59.5	66.1	642	US-10-369-493-21526	Sequence 21526, A
19	59.5	66.1	970	US-10-437-963-19810	Sequence 19810, A
20	59	65.6	890	US-10-437-963-118530	Sequence 118530, A
21	57.5	63.9	350	US-10-424-599-241211	Sequence 241211, A
22	57.5	63.9	473	US-10-425-114-43311	Sequence 43311, A
23	57.5	63.9	133	US-10-424-599-247410	Sequence 247410, A
24	57	63.3	312	US-10-425-114-45797	Sequence 45797, A
25	57	63.3	718	US-10-437-963-132395	Sequence 132395, A
26	57	63.3	1361	US-10-437-963-165703	Sequence 165703, A
27	55	61.1	237	US-10-424-599-202249	Sequence 202249, A
28	55	61.1	262	US-10-424-599-252174	Sequence 252174, A
29	53	58.9	256	US-10-437-963-111797	Sequence 111797, A
30	53	58.9	1944	US-10-369-493-2521	Sequence 2521, Ap
31	52.5	58.3	813	US-10-437-963-182704	Sequence 182704, A
32	52	57.8	712	US-10-369-493-22762	Sequence 22762, A
33	52	57.8	828	US-10-144-194A-96	Sequence 96, Appl
34	52	57.8	2677	US-10-144-194A-22	Sequence 22, Appl
35	48	53.3	830	US-10-369-493-4012	Sequence 4012, Ap
36	47.5	52.8	414	US-10-652-334-28	Sequence 28, Appl
37	47.5	52.8	683	US-10-369-493-22264	Sequence 22264, A
38	47	52.2	2646	US-10-437-963-189923	Sequence 189923, A
39	46.5	51.7	648	US-10-369-493-20334	Sequence 20334, A
40	46.5	51.7	650	US-10-369-493-2884	Sequence 2884, Ap
41	46.5	51.7	655	US-10-828-924-80	Sequence 80, Appl
42	46.5	51.7	656	US-10-369-493-1268	Sequence 1268, Ap
43	46	51.1	413	US-10-094-749-2124	Sequence 2124, Ap
44	46	51.1	3234	US-10-093-463-168	Sequence 168, Ap
45	45.5	50.6	653	US-10-369-493-21645	Sequence 21645, A

#### ALIGNMENTS

RESULT 1  
US-10-652-334-6  
Sequence 6, Application US/10652334  
Publication No. US20040115787A1  
GENERAL INFORMATION:  
APPLICANT: Pellet, Stuart  
APPLICANT: Cephalinski, Kevin  
APPLICANT: Dimmar, Jonathan D.  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
FILE REFERENCE: 601-1-85N  
CURRENT FILING DATE: 2003-08-28  
PRIOR APPLICATION NUMBER: US/10/652,334  
PRIOR FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
PRIOR FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 21  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
FEATURES:  
OTHER INFORMATION: Xaa = any amino acid  
US-10-652-334-6

Query Match 86.7%; Score 78; DB 16; Length 21;  
Best Local Similarity 100.0%; Pred. No. 3.3e-07;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 LXXQYRHPXISEFPYXGXL 21  
1 LXXQYRHPXISEFPYXGXL 21

RESULT 2  
US-10-652-334-25  
; Sequence 25, Application US/10652334  
; Publication No. US20040115787A1  
; GENERAL INFORMATION:  
; APPLICANT: Peltz, Stuart  
; APPLICANT: Czaplinski, Kevin  
; APPLICANT: Dinman, Jonathan D.  
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
; FILE REFERENCE: 601-1-85N  
; CURRENT APPLICATION NUMBER: US/10/652,334  
; PRIOR FILING DATE: 2003-08-28  
; PRIOR APPLICATION NUMBER: US/09/359,268A  
; PRIOR FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 60/093,685  
; PRIOR FILING DATE: 1998-07-22  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 25  
; LENGTH: 415  
; TYPE: PRT  
; ORGANISM: *Saccharomyces cerevisiae*  
US-10-652-334-25

Query Match 75.0%; Score 67.5; DB 16; Length 415;  
Best Local Similarity 62.5%; Pred. No. 0.00068;  
Matches 15; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

OY 1 LXXQYRMHPKISEPFX--YXGXL 21  
Db 213 LQVQYRMHPKISEPFIKKIYNGEL 236

RESULT 3  
US-10-767-701-38822  
; Sequence 38822, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 38822  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: *Sorghum bicolor*  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28WAY03-C7426\_1.pep  
US-10-767-701-38822

Query Match 72.8%; Score 65.5; DB 16; Length 199;  
Best Local Similarity 58.3%; Pred. No. 0.00071;  
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

OY 1 LXXQYRMHPKISEPFX--YXGXL 21  
Db 87 LQVQYRMHPKISEPNSCFYEGTL 110

RESULT 4  
US-10-424-599-205643  
; Sequence 205643, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 205643  
; LENGTH: 734  
; TYPE: PRT  
; ORGANISM: *Glycine max*  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1) - (734)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_27723C.1.pep  
US-10-424-599-205643

Query Match 72.8%; Score 65.5; DB 15; Length 734;  
Best Local Similarity 58.3%; Pred. No. 0.0029;  
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

OY 1 LXXQYRMHPKISEPFX--YXGXL 21  
Db 178 LQVQYRMHPKISEPNSFYEGTL 201

RESULT 5  
US-10-369-493-2087  
; Sequence 2087, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2087  
; LENGTH: 925  
; TYPE: PRT  
; ORGANISM: *Schizosaccharomyces pombe*  
US-10-369-493-2087

Query Match 72.8%; Score 65.5; DB 15; Length 925;  
Best Local Similarity 58.3%; Pred. No. 0.0037;  
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

OY 1 LXXQYRMHPKISEPFX--YXGXL 21  
Db 626 LQVQYRMHPKISEPNSFYEGTL 649

RESULT 6  
US-10-369-493-12790  
; Sequence 12790, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng



```

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12790
; LENGTH: 992
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
US-10-369-493-12790

Query Match      72.8%; Score 65.5; DB 15; Length 992;
Best Local Similarity 58.3%; Pred. No. 0.004;
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

Oy      1 LXXQYRMHPXISRPX---YXGXL 21
Db      566 LNVQYRMHPXISRPXSNMFTYGSL 589

RESULT 7
US-10-652-334-26
; Sequence 26, Application US/10652334
; Publication No. US20040115787A1
; GENERAL INFORMATION:
; APPLICANT: Pelcz, Stuart
; APPLICANT: Czaplinski, Kevin
; APPLICANT: Dلمان, Jonathan D.
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 601-1-85N
; CURRENT APPLICATION NUMBER: US/10/652,334
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/359,268A
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,685
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 472
; TYPE: PRT
; ORGANISM: saccharomyces cerevisiae
US-10-652-334-26

Query Match      71.7%; Score 64.5; DB 16; Length 472;
Best Local Similarity 58.3%; Pred. No. 0.0027;
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

Oy      1 LXXQYRMHPXISRPX---YXGXL 21
Db      301 LNVQYRMHPXISRPSSSEFYQRL 324

RESULT 8
US-10-369-493-1830
; Sequence 1830, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28

```

```

; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1830
; LENGTH: 2231
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1830

Query Match      71.7%; Score 64.5; DB 15; Length 2231;
Best Local Similarity 58.3%; Pred. No. 0.015;
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

Oy      1 LXXQYRMHPXISRPX---YXGXL 21
Db      1651 LNVQYRMHPXISRPSSSEFYQRL 1674

RESULT 9
US-10-767-701-49859
; Sequence 49859, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 49859
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3478-020-P1-K1-F11.pep
US-10-767-701-49859

Query Match      70.6%; Score 63.5; DB 16; Length 163;
Best Local Similarity 58.3%; Pred. No. 0.0013;
Matches 14; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

Oy      1 LXXQYRMHPXISRPX---YXGXL 21
Db      82 LNVQYRMHPXISRPSSSEFYQRL 105

RESULT 10
US-10-437-963-181690
; Sequence 181690, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 181690
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:

```

NAME/KEY: unsure  
LOCATION: (1)..(638)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_7893C.1.pap  
US-10-437-963-181690

Query Match 69.4%; Score 62.5; DB 16; Length 638;  
Best Local Similarity 54.2%; Pred. No. 0.0088;  
Matches 13; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 1 LXXQYRMHPXISEPPX---YXGXL 21  
DB 94 LQVQYRMHPCLSDPSPNCFYEGSL 117

RESULT 11  
US-10-029-386-31794  
Sequence 31794, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: AECOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 31794  
LENGTH: 71  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC003972.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.8  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3  
OTHER INFORMATION: SWISSPROT HIT: Q09820, EVALU2.00e-24  
US-10-029-386-31794

Query Match 68.3%; Score 61.5; DB 14; Length 71;  
Best Local Similarity 54.2%; Pred. No. 0.0013;  
Matches 13; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 1 LXXQYRMHPXISEPPX---YXGXL 21  
DB 42 LQVQYRMHPALSAFSPNIFYEGSL 65

RESULT 12  
US-10-408-765A-1706  
Sequence 1706, Application US/10408765A  
Publication No. US20040101874A1  
GENERAL INFORMATION:  
APPLICANT: Ghosh, Soumitra S.  
APPLICANT: Fahy, Eoin D.  
APPLICANT: Zhang, Bing  
APPLICANT: Gibson, Bradford W.  
APPLICANT: Taylor, Steven W.  
APPLICANT: Glenn, Gary M.  
APPLICANT: Watson, Dale E.  
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
FILE REFERENCE: 660088.465  
CURRENT APPLICATION NUMBER: US/10/408,765A  
CURRENT FILING DATE: 2003-04-04  
NUMBER OF SEQ ID NOS: 3077  
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1706  
LENGTH: 935  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-408-765A-1706

Query Match 68.3%; Score 61.5; DB 16; Length 935;  
Best Local Similarity 54.2%; Pred. No. 0.02;  
Matches 13; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 1 LXXQYRMHPXISEPPX---YXGXL 21  
DB 698 LQVQYRMHPALSAFSPNIFYEGSL 721

RESULT 13  
US-10-474-553-6  
Sequence 6, Application US/10474553  
Publication No. US20040161765A1  
GENERAL INFORMATION:  
APPLICANT: JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING DISEASE  
FILE REFERENCE: JHY-020.25  
CURRENT APPLICATION NUMBER: US/10/474,553  
CURRENT FILING DATE: 2003-10-10  
PRIOR APPLICATION NUMBER: 60/283,920  
PRIOR FILING DATE: 2001-04-13  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 1118  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-474-553-6

Query Match 68.3%; Score 61.5; DB 16; Length 1118;  
Best Local Similarity 54.2%; Pred. No. 0.025;  
Matches 13; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 1 LXXQYRMHPXISEPPX---YXGXL 21  
DB 698 LQVQYRMHPALSAFSPNIFYEGSL 721

RESULT 14  
US-10-767-701-57601  
Sequence 57601, Application US/10767701  
Publication No. US20040172684A1  
GENERAL INFORMATION:  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
FILE REFERENCE: 38-21(5335)B  
CURRENT APPLICATION NUMBER: US/10/767,701  
CURRENT FILING DATE: 2004-01-29  
NUMBER OF SEQ ID NOS: 63128  
SEQ ID NO 57601  
LENGTH: 219  
TYPE: PRT  
ORGANISM: Sorghum bicolor  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(219)  
OTHER INFORMATION: unsure at all Xaa locations  
OTHER INFORMATION: unsure at all Xaa locations  
OTHER INFORMATION: Clone ID: 30969562.pap  
US-10-767-701-57601

Query Match 67.2%; Score 60.5; DB 16; Length 219;  
Best Local Similarity 58.3%; Pred. No. 0.0065;

Matches 14; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

QY 1 LXXQYRMHPXISEPPX---YXGXL 21

Db 115 LKIQYRMHPXISEPPXKEPREGVL 138

## RESULT 15

US-10-437-963-140079

; Sequence 140079, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OR INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 140079

; LENGTH: 1975

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_41310C.1.pep

US-10-437-963-140079

Query Match 66.7%; Score 60; DB 16; Length 1975;

Best local similarity 73.3%; Pred. No. 0.085;

Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LXXQYRMHPXISEPP 15

Db 1452 LTKQYRMHPXISEPP 1466

Search completed: April 18, 2005, 09:04:07  
Job time : 75.4964 secs

This page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using bw model

Run on: April 18, 2005, 07:32:05 ; Search time 20.2336 Seconds  
(without alignments)  
99.861 Million cell updates/sec

Title: US-10-652-334-6  
Perfect score: 90  
Sequence: 1 LXXQYRMHPXISEPPYXGXL 21

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.79:.\*  
2: PIR1:.\*  
3: PIR2:.\*  
4: PIR3:.\*  
5: PIR4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67.5	75.0	1121	2	330862 DNA dependent ATPase
2	66.5	73.9	1069	2	713280 nonmense-mediated
3	65.5	72.8	935	2	562476 hypothetical prote
4	64.5	71.7	2231	2	553416 SEN1 protein - yea
5	62.5	69.4	1825	2	T52521 related to SEN1 pr
6	60.5	67.2	1090	2	T00533 probable DNA2-NAM7
7	60.5	67.2	1311	2	T08936 hypothetical prote
8	60	66.7	1687	2	T19072 DNA2-NAM7 helicase
9	59.5	66.1	942	2	D59085 transcrition cont
10	59.5	66.1	971	2	D53408 prematurely termin
11	59	65.6	692	2	E90113 hypothetical prote
12	57.5	63.9	660	2	P85069 hypothetical prote
13	57	63.3	1075	2	C96682 protein F1B22.16
14	55	61.1	2142	2	D6303 F17F16.1 protein
15	53	58.9	1944	2	T40065 tRNA-splicing endo
16	52	57.8	660	2	T41580 probable dna-bind
17	50.5	56.1	751	2	C84367 DNA binding protei
18	50	55.6	1076	2	B96682 protein F1B22.14
19	47.5	52.8	683	2	G34700 probable purine nu
20	46.5	51.7	650	2	G2429 hypothetical prote
21	46.5	51.1	368	1	E71080 probable DNA-bind
22	46	51.1	369	1	FOADM5 minor core protein
23	46	51.1	369	1	FOADM2 minor core protein
24	45.5	50.6	653	2	B75105 probable DNA helic
25	45	50.0	246	2	P86784 hypothetical prote
26	44	48.9	663	2	H64312 probable DNA helic
27	43.5	48.3	239	2	T64411 hypothetical prote
28	43.5	48.3	648	2	C69423 DNA helicase homol
29	43	47.8	662	2	B83201 conserved hypothe

30	41.5	46.1	1004	2	A39611 probable GTP-bind
31	41.5	46.1	1048	2	C86189 protein T25N20.11
32	41	45.6	555	2	E71420 hypothetical prote
33	40.5	45.0	989	2	T48845 insulin II gene en
34	40	44.4	389	2	S00629 pectinesterase (EC
35	40	44.4	515	2	I49154 calcitonin recepto
36	40	44.4	546	2	S46527 pectinesterase (EC
37	40	44.4	550	2	S46528 pectinesterase (EC
38	39	43.3	139	2	A83975 hypothetical prote
39	39	43.3	347	2	S33939 minor core protein
40	39	43.3	478	2	A37430 calcitonin recepto
41	39	43.3	479	2	S33746 calcitonin recepto
42	39	43.3	515	2	I60800 calcitonin recepto
43	38.5	42.8	530	2	D70476 DNA helicase - Agu
44	38	42.2	157	2	F81274 probable flagellar
45	38	42.2	283	2	F75265 conserved hypothe

#### ALIGNMENTS

RESULT 1  
S30862  
DNA dependent ATPase/DNA helicase B - yeast (Saccharomyces cerevisiae)  
N/Alternate names: protein YER176w  
C/Species: Saccharomyces cerevisiae  
C/Date: 28-May-1993 #sequence revision 28-May-1993 #text\_change 09-Jul-2004  
C/Accession: S30862; S50679; J02490; PC2368  
R/Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor, I  
submitted to the EMBL Data Library, February 1993  
A/Accession: S30862  
A/Molecule type: DNA  
A/Residues: 1-1121 <DIB>  
A/Cross-references: UNIPROT:P32644; GB:U18922; EMBL:L11229; NID:G603405; P1DN:AA64703.1  
R/Dietrich, F.S.  
submitted to the EMBL Data Library, December 1994  
A/Description: The sequence of S. cerevisiae cosmid 9163 and 9132.  
A/Reference number: S50679  
A/Accession: S50679  
A/Molecule type: DNA  
A/Residues: 1-1121 <DIB>  
A/Cross-references: EMBL:U18922; NID:G603405; P1DN:AA64703.1; P1D:G603417; MIPS:YER176w  
R/Blewas, B.E.; Chen, P.H.; Leezyk, J.; Blewas, S.B.  
Biochem. Biophys. Res. Commun. 206, 850-856, 1995  
A/Title: Biochemical and genetic characterization of a replication protein A dependent D  
A/Reference number: J02490; M01D:95134267; PMID:7832796  
A/Accession: J02490  
A/Molecule type: DNA  
A/Residues: 1-1121 <BIS>  
A/Accession: PC2368  
A/Molecule type: protein  
A/Residues: 277-283;623-633; 'X', 635-643 <BIS>  
C/Comment: This enzyme plays pivotal roles in the unwinding of the DNA double helix dur  
C/Genetics:  
A/Gene: SGD:EGM32  
A/Cross-references: SGD:S0000978; MIPS:YER176w  
A/Map position: 3K  
C/Keywords: nucleotide binding; P-loop  
F:670-677/Region: nucleotide-binding motif A (P-loop)

Query Match 75.0%; Score 67.5; DB 2; Length 1121;  
Best Local Similarity 62.5%; Pred. No. 0.00043;  
Matches 15; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

Cy 1 LXXQYRMHPXISEPP---YXGXL 21  
Db 880 LDTQYRMHPXISEPPKIKYNGEL 903

RESULT 2  
T43280  
nonsense-mediated mRNA decay trans-acting factor - Caenorhabditis elegans

C/Species: Caenorhabditis elegans  
C/Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C/Accession: T43280  
R/Page, M.F.; Carr, B.; Anders, K.R.; Grimson, A.; Anderson, P.  
Mol. Cell. Biol. 19, 5943-5951, 1999  
A/Title: SWG-2 is a phosphorylated protein required for mRNA surveillance in Caenorhabditis  
A/Reference number: 222389; MUID:99384262; PMID:10454541  
A/Accession: T43280  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-1069 <PAG>  
A/Cross-references: UNIPROT:O76512; EMBL:AF074017; NID:g3328176; PIDN:AAC26789.1; PID:g3328176

Query Match 73.9%; Score 66.5; DB 2; Length 1069;  
Best Local Similarity 58.3%; Pred. No. 0.00062;  
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

OY 1 LXXQYRMHPXISFPPX---YXGXL 21  
Db 672 LVQVYRMHPVLSEFPNMFYDGL 695

RESULT 3  
S62476  
hypothetical protein SPAC16C9.06c - fission yeast (Schizosaccharomyces pombe)  
C/Species: Schizosaccharomyces pombe  
C/Date: 16-May-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C/Accession: S62476; T37779  
R/Badcock, K.; Churcher, C.M.  
submitted to the EMBL Data Library, October 1995  
A/Reference number: S62445  
A/Accession: S62476  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-935 <BAD>  
A/Cross-references: UNIPROT:Q09820; EMBL:Z54366; NID:g1019812; PIDN:CAA91194.1; PID:g1324  
R/Badcock, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, October 1995  
A/Reference number: 221745  
A/Accession: T37779  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-935 <BA2>  
A/Cross-references: EMBL:Z54366; PIDN:CAA91194.2; GSPDB:GN00066; SPDB:SPAC16C9.06c  
A/Experimental source: strain 972h; cosmid c16C9  
C/Genetics:  
A/Gene: SPDB:SPAC16C9.06c  
A/Map position: 1L

Query Match 72.8%; Score 65.5; DB 2; Length 935;  
Best Local Similarity 58.3%; Pred. No. 0.00082;  
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

OY 1 LXXQYRMHPXISFPPX---YXGXL 21  
Db 636 LVQVYRMHPCLSEFPNMFYEGTL 659

RESULT 4  
S53416  
SEN1 protein - yeast (Saccharomyces cerevisiae)  
N/Alternate names: protein L9576.1; protein YLR430W  
C/Species: Saccharomyces cerevisiae  
C/Date: 05-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 09-Jul-2004  
C/Accession: S53416; A44387; S41985  
R/Favell, A.  
submitted to the EMBL Data Library, February 1995  
A/Description: The sequence of S. cerevisiae cosmid 9576.  
A/Reference number: S53409  
A/Accession: S53416  
A/Molecule type: DNA  
A/Residues: 1-2231 <PAV>  
A/Cross-references: UNIPROT:Q00416; EMBL:U20939; NID:g664871; PID:g664872; MIPS:YLR430W

A/Experimental source: strain S288C (AB972)  
R/dedharini, D.J.; Winey, M.; Ureic, D.; Webb, F.; Culbertson, M.R.  
Mol. Cell. Biol. 12, 2154-2164, 1992  
A/Title: SEN1, a positive effector of tRNA-splicing endonuclease in Saccharomyces cerevisiae  
A/Reference number: A44387; MUID:92236590; PMID:1569945  
A/Accession: A44387  
A/Molecule type: DNA  
A/Residues: 'MHS', 130, 'PCREVC', 131-2231 <DEM>  
A/Cross-references: GB:M74589; NID:g172573; PIDN:AAB63976.1; PID:g172574  
C/Genetics:  
A/Gene: SCD:SEN1  
A/Cross-references: SCD:S0004422; MIPS:YLR430W  
A/Map position: 12R  
C/Function:  
A/Description: may be component of nuclear splicing complex  
C/Keywords: nucleotide binding; nucleus; P-loop  
F1357-1364/Region: nucleotide-binding motif A (P-loop)

Query Match 71.7%; Score 64.5; DB 2; Length 2231;  
Best Local Similarity 58.3%; Pred. No. 0.0032;  
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

OY 1 LXXQYRMHPXISFPPX---YXGXL 21  
Db 1651 LDVQYRMHPISKEPSSRFYQGL 1674

RESULT 5  
T52521  
related to SEN1 protein [imported] - Neurospora crassa  
N/Alternate names: protein B2J23.170  
C/Species: Neurospora crassa  
C/Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004  
C/Accession: T52521  
R/Schulze, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fatmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, September 2000  
A/Reference number: 226053  
A/Accession: T52521  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-1825 <SCH>  
A/Cross-references: UNIPROT:Q9HFI5; EMBL:AL442164; GSPDB:GN00116; NCSP:B2J23.170  
A/Experimental source: BAC clone B2J23; strain OR74A  
C/Genetics:  
A/Gene: NCSP:B2J23.170  
A/Map position: 6

Query Match 69.4%; Score 62.5; DB 2; Length 1825;  
Best Local Similarity 58.3%; Pred. No. 0.006;  
Matches 14; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

OY 1 LXXQYRMHPXISFPPX---YXGXL 21  
Db 1465 LDVQYRMHPISFPPRAFYBGL 1488

RESULT 6  
T00533  
probable DNA2-NAM7 helicase family protein [imported] - Arabidopsis thaliana  
N/Alternate names: SEN1 protein homolog T20K24.14  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 09-Jul-2004  
C/Accession: T00533; G84572  
R/Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, July 1997  
A/Description: Arabidopsis thaliana chromosome II BAC T20K24 genomic sequence.  
A/Reference number: Z14167  
A/Accession: T00533  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1090 <ROU>  
A/Cross-references: UNIPROT:O64476; EMBL:AC002392; NID:g3176701; PID:g3176714  
A/Experimental source: cultivar Columbia

R.Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKern, S.E.; Unayam, L.; Tallon, L.; Eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402:761-768, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A/Reference number: A84420; MUID:20083487; PMID:10617197  
A/Accession: G84572

A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-1090 <STO>  
A/Cross-references: GB:AB020293; NID:g3176714; PIDN:AMD12029.1; GSPDB:GN00139  
C/Genetics:  
A/Gene: T20K24.14; AC2919120  
A/Map position: 2  
A/Intons: 250/1; 310/1; 342/1; 383/2; 427/3; 1016/3

Query Match 67.2%; Score 60.5; DB 2; Length 1090;  
Best Local Similarity 54.2%; Pred. No. 0.008;  
Matches 13; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 1 LXXQYRHPXISRPX---YXGX 21  
DB 843 LTVQYRHPQIRDPFSRYFGRL 866

## RESULT 7

hypothetical protein F6G3.130 - *Arabidopsis thaliana*  
C/Species: *Arabidopsis thaliana* (mouse-ear cress)

C/Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C/Accession: T08986

R/Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Meyer, K.F. X submitted to the Protein Sequence Database, May 1999

A/Reference number: Z16520  
A/Accession: T08986

A/Molecule type: DNA  
A/Residues: 1-1311 <BEV>

A/Cross-references: UNIPROT:Q9S2M3; EMBL:AL078464; GSPDB:GN00062; ATSP:F6G3.130  
A/Experimental source: cultivar Columbia; BAC clone F6G3

C/Genetics:  
A/Gene: ATSP:F6G3.130

A/Map position: 4  
A/Intons: 414/1; 474/1; 506/1; 547/2; 591/3; 1179/3

Query Match 67.2%; Score 60.5; DB 2; Length 1311;  
Best Local Similarity 54.2%; Pred. No. 0.0098;  
Matches 13; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 1 LXXQYRHPXISRPX---YXGX 21  
DB 1006 LTVQYRHPQIRDPFSRYFGRL 1029

## RESULT 8

DNM2-NM7 helicase family protein - fission yeast (*Schizosaccharomyces pombe*)  
C/Species: *Schizosaccharomyces pombe*

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: T39072

R/Murphy, L.; Harris, D.; Bartell, B.G.; Rajandream, M.A.; Connor, R.E.; Wood, V. submitted to the EMBL Data Library, August 1997

A/Reference number: Z21825  
A/Accession: T39072

A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA

A/Residues: 1-1687 <MUR>  
A/Cross-references: UNIPROT:Q92355; EMBL:Z81317; PIDN:CAB03612.1; GSPDB:GN00066; SPDB:SF

C/Genetics:  
A/Experimental source: strain 972h-; cosmid c6G9  
A/Gene: SPDB:SPAC69.010C

A/Map position: 1

Query Match 66.7%; Score 60; DB 2; Length 1687;

Best Local Similarity 73.3%; Pred. No. 0.016;  
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LXXQYRHPXISRP 15  
DB 1441 LS1QYRHPDISHP 1455

## RESULT 9

transcription control factor enhancer-binding protein - *Methanobacterium thermoautotroph*  
C/Species: *Methanobacterium thermoautotrophicum*

C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C/Accession: D69085

R/Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; E.; Liu, D.; Spatafore, R.; Viscare, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997  
A>Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: functi

A/Reference number: A69000; MUID:98037514; PMID:9371463  
A/Accession: D69085

A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA

A/Residues: 1-642 <MTH>

A/Cross-references: UNIPROT:O27671; GB:AE00922; GB:AE00666; NID:g2622754; PIDN:AA86107  
A/Experimental source: strain Delta H

C/Genetics:  
A/Gene: MTH1634

C/Superfamily: probable DNA helicase MJ0104

Query Match 66.1%; Score 59.5; DB 2; Length 642;  
Best Local Similarity 54.2%; Pred. No. 0.0069;  
Matches 13; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 1 LXXQYRHPXISRP---YXGX 21  
DB 429 LNCQYRHPAIMEPVRNRPYDGR 452

## RESULT 10

S23408  
prematurely terminated mRNA decay factor NAM7 - yeast (*Saccharomyces cerevisiae*)

N/Alternate names: protein YMS83.05c; protein YMR080c; UPF1 protein  
C/Species: *Saccharomyces cerevisiae*

C/Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
C/Accession: S23408; A44388; S54455

R/Altamir, N.; Groudinsky, O.; Dujardin, G.; Slonimski, P.P.  
J. Mol. Biol. 224, 575-587, 1992

A>Title: NAM7 nuclear gene encodes a novel member of a family of helicases with a Zn-11g  
A/Reference number: S23408; MUID:92235815; PMID:1314899

A/Accession: S23408  
A/Molecule type: DNA

A/Residues: 1-971 <ALT>

A/Cross-references: UNIPROT:P30771; EMBL:X62394; NID:g4022; PIDN:CAA44266.1; PID:g4023  
Mol. Cell. Biol. 12, 2165-2177, 1992

A/Title: Gene products that promote mRNA turnover in *Saccharomyces cerevisiae*.  
A/Reference number: A44388; MUID:92236591; PMID:1569946

A/Accession: A44388  
A/Molecule type: DNA

A/Residues: 1-971 <LEB>  
A/Cross-references: GB:M76659; NID:g173141; PIDN:AAA35197.1; PID:g173142

R/Gentles, S.; Bowman, S. submitted to the EMBL Data Library, May 1995

A/Reference number: S54455  
A/Accession: S54455

A/Molecule type: DNA  
A/Residues: 1-971 <GEN>  
A/Cross-references: EMBL:Z49259; NID:g807956; PID:g807962; MIPS:YMR080C

C/Genetics:  
A/Experimental source: strain AB972  
A/Gene: SGD:NAM7; UPF1

A/Cross-references: SGD:S0004685; MIPS:YMR080C

A:Map position: 13R  
 C:Keywords: GTP binding; mitochondrion; nucleotide binding; nucleus; P-loop  
 P:430-437/Region: nucleotide-binding motif A (P-loop)  
 P:545-548/Region: GTP-binding NKXD motif

Query Match 66.1%; Score 59.5; DB 2; Length 971;  
 Best Local Similarity 54.2%; Pred. No. 0.011;  
 Matches 13; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 1 LXXQYRHPXISEPPX---YXGX 21  
 DB 634 LEVQYRNPYLSEPPSNMFEGSL 657

RESULT 11  
 E90113  
 hypothetical protein component of a tRNA splicing complex [imported] - Gulliardia theta

C:Species: nucleomorph Gulliardia theta  
 A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
 C:Accession: E90113  
 R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rel  
 Nature 410, 1091-1096, 2001  
 A:Title: The highly reduced genome of an enslaved algal nucleus.  
 A:Reference number: A99082; MUID:11323671; PMID:11323671

A:Accession: E90113  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-692 <DOU>  
 A:Cross-references: UNIPROT:Q9AVZ7; GB:AF010592; NID:g12580756; PIDN:CAIC27074.1; GSPDB:G  
 C:Genetics:  
 A:Gene: component of a tRNA splicing complex  
 A:Map position: 2  
 A:Genome: nucleomorph  
 C:Keywords: nucleomorph

Query Match 65.6%; Score 59; DB 2; Length 692;  
 Best Local Similarity 73.3%; Pred. No. 0.0092;  
 Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LXXQYRHPXISEPP 15  
 DB 451 LGIQYRHPQISPP 465

## RESULT 12

F85069  
 hypothetical protein AT4G05540 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: F85069  
 R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
 Nature 402, 769-777, 1999  
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
 A:Reference number: A85001; MUID:20083488; PMID:10617198

A:Accession: F85069  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-660 <STO>  
 A:Cross-references: UNIPROT:Q9S9W3; GB:NC\_001268; NID:g7267314; PIDN:CAB81096.1; GSPDB:G  
 C:Genetics:  
 A:Gene: AT4G05540  
 A:Map position: 4

Query Match 63.9%; Score 57.5; DB 2; Length 660;  
 Best Local Similarity 54.2%; Pred. No. 0.017;  
 Matches 13; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 1 LXXQYRHPXISEPP---YXGX 21  
 DB 471 LNVQYRHPISISLPNMFYGGKI 494

## RESULT 13

C96682  
 protein F1R22.16 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: C96682  
 R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Hutzar, L.  
 Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, I  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: C96682  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1075 <STO>  
 A:Cross-references: UNIPROT:Q9SHX6; GB:AE005173; NID:g6686402; PIDN:AF23836.1; GSPDB:GNC  
 C:Genetics:  
 A:Gene: F1R22.16  
 A:Map position: 1

Query Match 63.3%; Score 57; DB 2; Length 1075;  
 Best Local Similarity 73.3%; Pred. No. 0.035;  
 Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LXXQYRHPXISEPP 15  
 DB 647 LNVQYRHPISISLP 661

## RESULT 14

D86303  
 F17F16.1 protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: D86303  
 R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Hutzar, L.  
 Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, I  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D86303  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-2142 <STO>  
 A:Cross-references: UNIPROT:Q9FWR3; GB:AE005172; NID:g954728; PIDN:AG09081.1; GSPDB:GNC  
 C:Genetics:  
 A:Gene: F17F16.1  
 A:Map position: 1

Query Match 61.1%; Score 55; DB 2; Length 2142;  
 Best Local Similarity 66.7%; Pred. No. 0.17;  
 Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 LXXQYRHPXISEPP 15  
 DB 1594 LVCQYRHPICRFP 1608

## RESULT 15

T40065  
 tRNA-splicing endonuclease positive effector - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe



C:Date: 03-Dec-1999 #sequence\_reviation 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T40065  
R:Lyle, M.; Rajandream, M.A.; Barrell, B.G.; Leleure, V.; Galibert, F.  
submitted to the EMBL Data Library, December 1998  
A:Reference number: Z1903  
A:Accession: T40065  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1944 <LYN>  
A:Cross-references: UNIPROT:O94387; EMBL:AL034463; PIDD:CAM2438.1; GSPDB:GN00067; SPDB  
A:Experimental source: strain 972h-; cosmid c29A10  
C:Genetic8:  
A:Gene: SPDB:SPBC29A10.10c  
A:Map position: 2

This Page Blank (uspto)



RA PubMed=12368865; DOI=10.1038/nature01099;  
 RA Carlton J.M., Angiolini S.V., Suh B.B., Koo J.T.W., Pertea M.,  
 RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,  
 RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bigwell S.L.,  
 RA Shalimov S.J., van Aken S.B., Riedmuller S.B., Feldblyum T.V.,  
 RA Cho J.K., Quackenbush J., Sedegah M., Shoabdi A., Cummings L.M.,  
 RA Florens L., Yates F.R. III, Raine J.D., Sinden R.B., Harris M.A.,  
 RA Cunningham D.A., Preiser F.R., Bergman L.W., Valdivia A.B.,  
 RA van Lin L.H., Unse C.J., Waters A.P., Smith H.O., White O.R.,  
 RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,  
 RA Carucci D.J.;  
 RT "Genome sequence and comparative analysis of the model rodent malaria  
 RT parasite Plasmodium yoelii yoelii.";  
 RL Nature 419:512-519(2002).  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL: AAB10100336; EAA20580.1; -  
 SQ SEQUENCE 1297 AA; 149791 MW; 4933BF9327D412C1 CRC64;

Query Match 75.6%; Score 68; DB 2; Length 1297;  
 Best Local Similarity 70.6%; Pred. No. 0.002;  
 Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LXXOYRHPXISRPXY 17  
 | ||||| : |||||  
 Db 896 LVOYRHPXISRPXY 912

RESULT 3  
 06C803 PRELIMINARY; PRT; 964 AA.  
 AC 06C803;  
 DT 25-OCT-2004 (TREMBlrel. 28. Created)  
 DT 25-OCT-2004 (TREMBlrel. 28. Last sequence update)  
 DT 25-OCT-2004 (TREMBlrel. 28. Last annotation update)  
 DE Yarrowia lipolytica chromosome D of strain CLIB99 of Yarrowia  
 DE lipolytica.  
 GN ORFNames=YAL10D23881g;  
 OS Yarrowia lipolytica CLIB99.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Dipodascaceae; Yarrowia.  
 CX NCBI\_TaxID=284591;  
 RN 1;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIB99;  
 RG Genolevures;  
 RA Lafontaine I., Sherman D., Fischer G., Durrens P., Casaregola S.,  
 RA Goffard N., Frangeul L., Aligle M., Anthouard V., Babour A., Barbe V.,  
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boissarie A., Boyer J., Catolico L., Confarotier F., de Daruvar A.,  
 RA Despons L., Fabre E., Faltreud C., Ferry-Dumazet H., Gropi A.,  
 RA Hantziye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicard J.M., Nikoleki M., Ozias S., Ozler-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
 RA Swenne D., Tekala F., Wesolowski-Louvel M., Weschof E., Wirth B.,  
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukhazha M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissbach J.,  
 RA Wincker P., Souciet J.L.;  
 RT "Genome evolution in yeasts.";  
 RL Nature 430:35-44(2004).  
 RN 12;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIB99;  
 RA Genoscope;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: CR382130; CAG81409.1; -  
 DR GO: GO:0000166; F:nucleotide binding; IEA.  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR011545; DEAD/DEAH\_N.  
 DR SMART: SM00382; AAA, 1.

DR SMART: SM00487; DEXDC; 1.  
 KW ATP-binding.  
 SQ SEQUENCE 964 AA; 107764 MW; 68C05A712597B8BD CRC64;

Query Match 75.0%; Score 67.5; DB 2; Length 964;  
 Best Local Similarity 58.3%; Pred. No. 0.0018;  
 Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

Qy 1 LXXOYRHPXISRPXY---YXGX 21  
 | ||||| : |||||  
 Db 650 LVOYRHPXISRPXYMTYEGSL 673

RESULT 4  
 YE06\_YEAST STANDARD; PRT; 1121 AA.  
 ID YE06\_YEAST  
 AC P32644;  
 DT 01-OCT-1993 (Rel. 27. Created)  
 DT 01-OCT-1993 (Rel. 27. Last sequence update)  
 DT 05-JUL-2004 (Rel. 44. Last annotation update)  
 DE Hypothetical 127.0 kDa protein in RAD24-BMH1 intergenic region.  
 GN OrderedLocustNames=YER176W; ORFNames=SYGP-ORF61;  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 CX NCBI\_TaxID=4932;  
 RN 1;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RC MEDLINE=97313264; PubMed=9169868;  
 RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,  
 RA Araujo R., Aviles B., Berno A., Brennan T., Carpenter J., Chen E.,  
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,  
 RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,  
 RA Lin D., Moseale D., Nakahara K., Nemach A., Norgren R., Osifner P.,  
 RA Oh C., Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,  
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";  
 RL Nature 387:78-81(1997).  
 CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U18922; AAB64703.1; -  
 CC PIR: S30862; S30862.  
 DR Germonline; 139253; -  
 DR SGD; S000000978; ECM32.  
 DR GO: GO:0005844; C:Polyome; IDA.  
 DR GO: GO:0003678; F:DNA helicase activity; IDA.  
 DR GO: GO:0006449; P:regulation of translational termination; IMP.  
 DR InterPro: IPR001410; DEAD.  
 DR SMART: SM00487; DEXDC; 1.  
 KW ATP-binding; Helicase; Hypothetical protein.  
 KW NP\_BIND 670 677 ATP (Potential).  
 FT SEQUENCE 1121 AA; 126970 MW; 641CA6810282A0 CRC64;

Query Match 75.0%; Score 67.5; DB 1; Length 1121;  
 Best Local Similarity 62.5%; Pred. No. 0.0021;  
 Matches 15; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

Qy 1 LXXOYRHPXISRP---YXGX 21  
 | ||||| : |||||  
 Db 880 LVOYRHPXISRPXIKKINGEL 903

RESULT 5  
 RNT1\_ARATH

```

ID  RNT1_ARATH STANDARD; PRT; 1235 AA.
AC  09FUR0;
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DE  05-JUL-2004 (Rel. 44, Last annotation update)
GN  Regulator of nonsense transcripts 1 homolog.
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC  eurosids II; Brassicales; Brassicaceae; Arabidopsids.
CX  NCBI_TaxID=3702;
[1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=cv Columbia;
RX  MEDLINE=96403884; PubMed=9734815;
RA  Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA  Tabata S.;
RT  "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT  Sequence features of the regions of 1,367,185 bp covered by 19
RT  physically assigned P1 and TAC clones."
RL  DNA Res. 5:203-216(1998).
CC  -1- FUNCTION: Eliminates the production of nonsense-containing RNAs
CC  (by similarity).
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC  -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
DR  EMBL; AB013394; BAB10240.1; -
DR  InterPro; IPR006935; Resili.
KW  ATP-binding; Helicase; Hydrolase; Hypothetical protein; Zinc-finger.
FT  ZN_FING 148 178 C2H2-type (atypical) (Potential).
FT  ZN_FING 202 232 C4-type (Potential).
FT  NP_BIND 507 514 ATP (Potential).
FT  DOMAIN 102 105 Poly-Ser.
SQ  SEQUENCE 1235 AA; 134870 MW; 177B3FC524D26A CRC64;
Query Match 75.0%; Score 67.5; DB 1; Length 1235;
Best Local Similarity 58.3%; Pred. No. 0.0024;
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;
QY 1 LXXQYRMHPXISFPPX---YKGL 21
Db 715 LQVQYRMHPALSFPPNSFYEGTL 738
RESULT 6
08S3K7 PRELIMINARY; PRT; 1243 AA.
AC  08S3K7;
DT  01-JUN-2002 (TReMBLrel. 21, Created)
DT  01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT  01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE  UP1 (Fragment).
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC  eurosids II; Brassicales; Brassicaceae; Arabidopsids.
CX  NCBI_TaxID=3702;
[1]
RN  SEQUENCE FROM N.A.
RP  Jorizuka Y., Mulligan M.R.;
RL  Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF484122; AAL92018.1; -
DR  GO; GO:0015668; P: type III site-specific deoxyribonuclease ac. .; IEA.
DR  GO; GO:0009307; P: DNA restriction; IEA.

```

```

DR  InterPro; IPR006935; Resili.
DR  Pfam; PF04851; Resili; 1.
FT  NON_TER 1
SQ  SEQUENCE 1243 AA; 135599 MW; 86920C9D0B3F7A86 CRC64;
Query Match 75.0%; Score 67.5; DB 2; Length 1243;
Best Local Similarity 58.3%; Pred. No. 0.0024;
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;
QY 1 LXXQYRMHPXISFPPX---YKGL 21
Db 707 LQVQYRMHPALSFPPNSFYEGTL 730
RESULT 7
RNT1_CAEEL STANDARD; PRT; 1069 AA.
AC  076512; Q9BL16;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  25-OCT-2004 (Rel. 45, Last annotation update)
DE  Regulator of nonsense transcripts 1 (Nonsense mRNA reducing factor 1)
DE  (Up-frameshift suppressor 1 homolog).
GN  Name=smg-2; ORFNames=Y48G9AL.6;
OS  Caenorhabditis elegans.
OC  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC  Rhabditidae; Peloderiinae; Caenorhabditis.
CX  NCBI_TaxID=6239;
[1]
RN  SEQUENCE FROM N.A. AND MUTAGENESIS OF GLY-470 AND GLY-472.
RP  MEDLINE=99384262; PubMed=10454541;
RA  Page M.F., Carr B., Anders K.R., Grierson A., Anderson P.;
RT  "SMG-2 is a phosphorylated protein required for mRNA surveillance in
RT  Caenorhabditis elegans and related to Upflp of yeast."
RL  Mol. Cell. Biol. 19:5943-5951(1999).
[2]
RN  SEQUENCE FROM N.A.
RP  STRAIN=Bristol N2;
RX  MEDLINE=99069613; PubMed=9851916;
RG  The C. elegans sequencing consortium;
RT  Investigating biology";
RT  "Genome sequence of the nematode C. elegans: a platform for
RL  Science 282:2012-2018(1998).
CC  -1- FUNCTION: Eliminates the production of nonsense-containing RNAs.
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC  -1- PTM: Phosphorylated. Smg-1, smg-3 and smg-4 are required for
CC  phosphorylation.
CC  -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
DR  EMBL; AF074017; AAC26789.1; -
DR  EMBL; AC025721; AAC29903.2; -
DR  PIR; T43280; T43280.
DR  WormBase; WBGene00004880; smg-2.
DR  WormBase; Y48G9AL.6; CB28367.
DR  InterPro; IPR003593; AAA ATPase.
DR  InterPro; IPR006935; Resili.
DR  SMART; SM00382; AAA; 1.
KW  ATP-binding; Helicase; Hydrolase; Phosphorylation; Zinc-finger.
FT  ZN_FING 103 131 C2H2-type (atypical) (Potential).
FT  ZN_FING 155 185 C4-type (Potential).
FT  NP_BIND 467 474 ATP (Potential).
FT  DOMAIN 991 1069 Gln/Ser-rich.
FT  DOMAIN 1029 1034 Poly-Thr.
FT  DOMAIN 1036 1039 Poly-His.
FT  MUTAGEN 470 470 G->R: In R866; loss of activity and

```

```

FT          472      472      Increased phosphorylation.
FT MUTAGEN      86      87      G->E; in R895; loss of activity and
FT CONFLICT      86      87      Increased phosphorylation.
FT CONFLICT      769      769      HE -> QO (in Ref. 2).
FT SEQUENCE      1069 AA; 120019 MW; 5B4C91FB49B5C24B CRC64;
SQ
Query Match      73.9%; Score 66.5; DB 1; Length 1069;
Best Local Similarity 58.3%; Pred. No. 0.0031;
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

Cy          1 LXXQYRMHPXISEPPX---YXGX 21
Db          672 LVQYRMHPVISEPPSNFYDGL 695

RESULT 8
ID Q6EFK3 PRELIMINARY; PRT; 1125 AA.
AC Q6EFK3;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
GN Candida glabrata strain CBS138 chromosome 1 complete sequence.
OS ORFNames=CAG10L12034g;
OC Candida glabrata CBS138.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitospotic Saccharomycetales; Candida.
OX NCBI_TaxID=284593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG Genoluvures;
RA Lafontaine I., de Montigny J., Marck C., Nevegliese C., Talia E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barde V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne B., Bleykasten C.,
RA Boismare A., Boyer J., Catolico L., Confiantieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerest A., Kovul R., Lemaitre M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolaki M., Oztas S., Ozler-Kalogeropoulos O.,
RA Pellenz S., Potter S., Richard G.F., Straub M.L., Suleau A.,
RA Sennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenoun-Meyer M., Zivanovic I., Bolotin-Fukhaz M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Soucier J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
DR EMBL; CR380958; CAC62265.1; -;
SQ SEQUENCE 1125 AA; 127137 MW; 63C0428123F3CC8C CRC64;

Query Match      73.3%; Score 66; DB 2; Length 1125;
Best Local Similarity 80.0%; Pred. No. 0.0041;
Matches 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy          1 LXXQYRMHPXISEPP 15
Db          885 LNTQYRMHPXISEPP 899

RESULT 9
RNTL SCHPO STANDARD; PRT; 925 AA.
AC Q09820;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Regulator of nonsense transcripts 1 homolog.
GN ORFNames=SPAC16C9.06c;
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetales;

```

```

OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voicakeert G., Aert R., Robben J., Grympez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moescl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wandut R., Purnelle B.,
RA Goffeau A., Cadieu S., Driano S., Gloux S., Lelaure V., Motier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Huxst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerniti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Sipakovski G.V., Usery D., Barrett B.G., Nure P.;
RL "The genome sequence of Schizosaccharomyces pombe.";
RT Nature 415:871-880(2002).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z54366; CAA91194.2; -;
DR PIR; S62476; S62476.
DR GeneDB; Spombe; SPAC16C9.06c; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR006935; ResIII.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00487; DEXDC; 1.
DR ATP-binding; Helicase; Hypothetical protein;
KW Nonsense-mediated mRNA decay; Zinc-finger.
FT ZN_FING 52 80 C2H2-type (atypical) (Potential).
FT ZN_FING 104 134 C4-type (Potential).
FT NP_BIND 414 421 ATP (By similarity).
SQ SEQUENCE 925 AA; 104528 MW; 4A5D63C82826E864 CRC64;

Query Match      72.8%; Score 65.5; DB 1; Length 925;
Best Local Similarity 58.3%; Pred. No. 0.0041;
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

Cy          1 LXXQYRMHPXISEPPX---YXGX 21
Db          626 LVQYRMHPXISEPPSNFYDGL 649

RESULT 10
ID Q8TFW3 PRELIMINARY; PRT; 1060 AA.
AC Q8TFW3;
DT 01-JUN-2002 (TRENBLrel. 21, Created)

```

DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)  
 DE Possible regulator of nonsense transcripts.  
 GN Name=AF5511.22c;  
 OS Apeeryllus fumigatus (Sartorya fumigata).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Apeeryllus.  
 NCBI\_Taxid=5085;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AF293;  
 RA Knowles D.G., Warren T., Hall N., Quail M., Woodward J.R.,  
 RA Denning D.W., Anderson M.J., Barrett B.  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL713629; CAD28448.1; -;  
 DR GO: GO:0005524; P:ATP binding; IEA.  
 DR GO: GO:0000166; P:nucleotide binding; IEA.  
 DR GO: GO:0015668; P:type III site-specific deoxyribonuclease ac. .; IEA.  
 DR GO: GO:0009307; P:DNA restriction; IEA.  
 DR InterPro: IPR003593; AAA ATPase.  
 DR InterPro: IPR011545; DEAD/DEAH\_N.  
 DR Pfam: PF04851; Resili; 1.  
 DR SMART: SM00382; AAA; 1.  
 DR SMART: SM00487; DEXDC; 1.  
 KW ATP-binding.  
 SQ SEQUENCE 1060 AA; 116728 MW; 744DFC58A26E77B CRC64;

Query Match 72.8%; Score 65.5; DB 2; Length 1060;  
 Best Local Similarity 58.3%; Pred. No. 0.0047;  
 Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

Qy 1 LXXQYRMHPXISEPPX---YXGXL 21  
 Db 634 LNVQYRMHPXCLSEPPSNMFYEGSL 657

RESULT 11  
 ID 06MY12 PRELIMINARY; PRT; 1079 AA.  
 AC 06MY12  
 DT 05-JUL-2004 (TREMblrel. 27, Created)  
 DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMblrel. 27, Last annotation update)  
 DE Regulator of nonsense transcripts, putative.  
 GN ORFNames=AF5511.22c;  
 OS Apeeryllus fumigatus (Sartorya fumigata).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Apeeryllus.  
 NCBI\_Taxid=5085;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC PubMed=1498527; DOI=10.1016/j.fgb.2003.12.003;  
 RA Pain A., Woodward J., Quail M.A., Anderson M.J., Clark R., Collins M.,  
 RA Foster N., Fraser A., Harris D., Larke N., Murphy L., Humphray S.,  
 RA O'Neill S., Pearce M., Price C., Rabbittsch E., Rajandream M.A.,  
 RA Salzberg S., Saunders D., Seegar K., Sharp S., Warren T.,  
 RA Denning D.W., Barrett B., Hall N.;  
 RA "Insight into the genome of Apeeryllus fumigatus: analysis of a 922  
 RT kb region encompassing the nitrate assimilation gene cluster."  
 RL Fungal Genet. Biol. 41:443-453 (2004).  
 DR EMBL: BX649506; CAP72021.1; -;  
 DR InterPro: IPR011545; DEAD/DEAH\_N.  
 DR InterPro: IPR005935; AAA ATPase.  
 DR Pfam: PF04851; Resili; 1.  
 DR SMART: SM00487; DEXDC; 1.  
 SQ SEQUENCE 1079 AA; 118655 MW; A8907BE6FA850063 CRC64;

Query Match 72.8%; Score 65.5; DB 2; Length 1079;  
 Best Local Similarity 58.3%; Pred. No. 0.0048;  
 Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;  
 Qy 1 LXXQYRMHPXISEPPX---YXGXL 21

Db 653 LNVQYRMHPXCLSEPPSNMFYEGSL 676

RESULT 12  
 ID RNT1\_NEUCR STANDARD; PRT; 1093 AA.  
 AC 09HEH1; Q7RVU9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Regulator of nonsense transcripts 1 homolog.  
 GN ORFNames=2E4.130; NCU04242.1;  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 NCBI\_Taxid=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=74-OR23-1A / FGSC 987;  
 RC MEDLINE=22542210; PubMed=12655011; DOI=10.1093/nar/gkg293;  
 RA Mannheim G., Montreux C., Haese D., Mewes H.-W., Algn V.,  
 RA Hohnel J.D., Faltmann B., Nyakatura G., Kempen F., Maier J.,  
 RA Schulte U.;  
 RA "What's in the genome of a filamentous fungus? Analysis of the  
 RT Neurospora genome sequence."  
 RL Nucleic Acids Res. 31:1944-1954 (2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=74-OR23-1A / FGSC 987;  
 RC PubMed=12712197; DOI=10.1038/nature01554;  
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
 RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehm B.,  
 RA Elting T., Engels R., Wang S., Nielsen C.B., Butler J., Endlitz M.,  
 RA Qui D., Janakiev P., Bell-Pedersen D., Nelson M.A., Braun E.L.,  
 RA Werner-Washburne M., Selitrenikoff C.P., Kinsey J.A., Braun E.L.,  
 RA Zelter A., Schulte U., Koche G.O., Jedd G., Mewes H.-W., Staben C.,  
 RA Marcotte E., Greenberg D., Roy A., Foley K., Naylor J.,  
 RA Stange-Thomann N., Barrett R., Gnerre S., Kamal M., Kamysaslis M.,  
 RA Mauceli E., Bielke C., Rudd S., Friedman D., Kryzofova S.,  
 RA Rasmussen C., Metzgerberg R.L., Perkins D.J., Kroken S., Cogoni C.,  
 RA Macino G., Catchside D., Li W., Pratt R.J., Osmati S.A.,  
 RA Desouza C.P., Glase L., Orbach M.J., Berglund J.A., Voelker R.,  
 RA Varden O., Plamann M., Sells S., Dunlap J., Radford A., Aramayo R.,  
 RA Nativig D.O., Alex L.A., Mannheim G., Ebbols D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nisbaum C., Birren B.,  
 RT "The genome sequence of the filamentous fungus Neurospora crassa."  
 RL Nature 422:859-868 (2003).  
 CC -1- FUNCTION: Eluminates the production of nonsense-containing RNAs  
 (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL: AL451022; CAC18314.1; -;  
 DR EMBL: AABK01000272; BA31397.1; -;  
 DR InterPro: IPR005593; AAA ATPase.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR005935; Resili.  
 DR SMART: SM00382; AAA; 1.  
 DR SMART: SM00487; DEXDC; 1.  
 KW ATP-binding; Helicase; Hypothetical protein;  
 KW Nonsense-mediated mRNA decay; Zinc-finger.  
 FT ZN\_FING 111 139 C2H2-type (atypical) (Potential).  
 FT FT 163 193 C4-type (Potential).  
 FT NP\_BIND 477 484 ATP (Potential).  
 FT

```

FT DOMAIN 59 62 Poly-Asp.
FT DOMAIN 69 73 Poly-Asp.
SQ SEQUENCE 1093 AA; 120087 MW; 880E4F0407ACE142 CRC64;

Query Match 72.8%; Score 65.5; DB 1; Length 1093;
Best Local Similarity 58.3%; Pred. No. 0.0049;
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 1 LXXQYRHPXISEFPX---YXGX 21
DB 683 LKQYRHPXCLSEFPXNFTGSL 706

RESULT 13
ID Q6BNH2 PRELIMINARY; PRT; 1124 AA.
AC Q6BNH2;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Debaryomyces hansenii chromosome E of strain CBS767 of Debaryomyces
hansenii.
GN ORFNames=DEHA0E23034g;
OS Debaryomyces hansenii CBS767.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxId=284592;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genolevures;
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Geoffard N., Frangoul L., Aigle M., Anthouard V., Babour A., Barde V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne B., Bleykasten C.,
RA Boisrame A., Boyer J., Caticolico L., Confiantolero F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Giropi A.,
RA Hantayre F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicoud J.M., Nikolaki M., Oztes S., Ozier-Kalogeropoulos O.,
RA Pellenné S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Svenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolocin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Soulier J.L.;
RA "Genome evolution in yeasts.";
RT Nature 430:35-44(2004).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=CBS767;
RC Genolevures;
RA Genoscope;
RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382137; F:GAG89524.1; -
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR001547; Glyco_hydro_5.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
KW ATP-binding.
SQ SEQUENCE 1124 AA; 127842 MW; D6F8B09E5614EE9 CRC64;

Query Match 72.2%; Score 65; DB 2; Length 1124;
Best Local Similarity 80.0%; Pred. No. 0.0062;
Matches 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LXXQYRHPXISEFP 15
DB 889 LDTQYRHPXISEFP 903

RESULT 14
ID Q7LIE9 PRELIMINARY; PRT; 757 AA.

```

```

AC Q7LIE9;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Snp1p (Fragment).
GN Name=SEN1;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxId=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97313267; Pubmed=9169871;
RA Johnston M., Hillier L., Riles L., Albertmann K., Andre B., Anorge W.,
RA Benes V., Bruckner M., Delius H., Dubois E., Dusterhoft A.,
RA Burtian K.D., Floeth M., Goffeau A., Hedling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Kotter P.,
RA Louis E.J., Messenguy F., Mewes H.W., Miosga T., Mostl D.,
RA Muller-Auer S., Nentwich U., Obermayer B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharte M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhaeselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hami J., Hohnsbeil J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Du Z.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Waterson K.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Cherry J.M.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U21094; AAB67523.1; -
DR SGD; S000004422; SEN1.
FT NON TER
SQ SEQUENCE 757 AA; 84887 MW; DA0B167986585B77 CRC64;

Query Match 71.7%; Score 64.5; DB 2; Length 757;
Best Local Similarity 58.3%; Pred. No. 0.005;
Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 1 LXXQYRHPXISEFPX---YXGX 21
DB 177 LKQYRHPXISKEPSEFTQGR 200

RESULT 15
ID Q6FKS5 PRELIMINARY; PRT; 1979 AA.
AC Q6FKS5;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Candida glabrata strain CBS138 chromosome M complete sequence.
GN ORFNames=CAG10M0990g;
OS Candida glabrata CBS138.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxId=284593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG Genolevures;

```



RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marck C., Neveuglise C., Talla B.,  
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Barney S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boistrame A., Boyer J., Cactolico L., Confanioleri P., de Daruvar A.,  
 RA Despons L., Fabre E., Falthead C., Ferry-Dumazet H., Groppi A.,  
 RA Hantreaye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Micaud J.M., Nikoleki M., Ozias S., Ozler-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
 RA Svenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
 RA Zenilou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissbach J.,  
 RA Wincker P., Souciet J.L.;  
 RT "Genome evolution in yeasts."  
 RL Nature 430:35-44(2004)  
 DR EMBL, CR380959; CAG62365.1; -;  
 SQ SEQUENCE 1979 AA; 224955 MW; A38P91B1F8E7CDFE CRC64;

Query Match 71.7%; Score 64.5; DB 2; Length 1979;  
 Best Local Similarity 58.3%; Pred. No. 0.014; 6; Indels 3; Gaps 1;  
 Matches 14; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 1 LXXQYRMHPXISEFPX---YXGXL 21  
 Db 1610 LDVQYRMHPAISKFPSSAFYDGR 1633

Search completed: April 18, 2005, 08:15:46  
 Job time : 88.6788 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: April 18, 2005, 07:10:15 ; Search time 69.6934 Seconds  
(without alignments)  
77.652 Million cell updates/sec

Title: US-10-652-334-7  
Perfect score: 54  
Sequence: 1 IGVTPTXXQVXXL 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues  
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneeqp1980s:\*  
2: geneeqp2000s:\*  
3: geneeqp2000s:\*  
4: geneeqp2000s:\*  
5: geneeqp2002s:\*  
6: geneeqp2003as:\*  
7: geneeqp2003bs:\*  
8: geneeqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	47	87.0	171 3 AAB33009	Aab33009 Pinus rad
2	46	85.2	14 7 ABW01183	Abw01183 Saccharom
3	46	85.2	14 8 ADP44109	Adp44109 Yeast tra
4	46	85.2	15 3 AAY77810	Aay77810 Motic VIT
5	46	85.2	648 8 ADS42927	Ads42927 Bacterial
6	46	85.2	1944 8 ADN19868	Adn19868 Bacterial
7	45	83.3	1274 4 ABB65781	Abb65781 Drosophi1
8	44	81.5	125 7 ADB63936	Adb63936 Human pro
9	44	81.5	414 3 AAY77816	Aay77816 Yeast Dip
10	44	81.5	414 7 ABW01204	Abw01204 Saccharom
11	44	81.5	414 8 ADP44130	Adp44130 Yeast hel
12	44	81.5	517 6 ABB99784	Abb99784 Amino aci
13	44	81.5	550 6 ABB99781	Abb99781 Amino aci
14	44	81.5	683 8 ADS43834	Ads43834 Bacterial
15	44	81.5	804 7 ADB65007	Adb65007 Human pro
16	44	81.5	1002 5 ABB79913	Abb79913 Arabidops
17	44	81.5	1187 4 AAU07865	Aau07865 Polypept1
18	44	81.5	1208 6 ABB99782	Abb99782 Amino aci
19	44	81.5	1211 4 AAU07886	Aau07886 Polypept1
20	44	81.5	1315 6 ABJ19762	Abj19762 Human MP2
21	44	81.5	1349 6 ABB99783	Abb99783 Amino aci
22	44	81.5	1942 7 ADES5894	Ades5894 Human pro
23	43	79.6	338 6 ABJ19763	Abj19763 Human MP2
24	43	79.6	642 8 ADS43096	Ads43096 Bacterial
25	43	79.6	650 8 ADN20231	Adn20231 Bacterial

26	42	77.8	415 3 AAY77813	Aay77813 Yeast Mtc
27	42	77.8	415 7 ABW01201	Abw01201 Saccharom
28	42	77.8	415 8 ADP44127	Adp44127 Yeast hel
29	42	77.8	1176 7 ABW02362	Abw02362 Klebslell
30	41	75.9	10 7 ABW01187	Abw01187 Saccharom
31	41	75.9	10 8 ADP44113	Adp44113 Yeast tra
32	41	75.9	578 4 ABG18338	Abg18338 Novel hum
33	41	75.9	1693 4 ABG17826	Abg17826 Novel hum
34	41	75.9	1693 4 ABG29081	Abg29081 Novel hum
35	41	75.9	2061 4 ABW1759	Abw1759 Drosophi1
36	40	74.1	50 4 AAM19141	Aam19141 Peptide #
37	40	74.1	50 4 ABB38350	Abb38350 Peptide #
38	40	74.1	50 4 AAM31789	Aam31789 Peptide #
39	40	74.1	50 4 ABB23527	Abb23527 Protein #
40	40	74.1	50 4 AAM71489	Aam71489 Human bon
41	40	74.1	50 4 AAM58962	Aam58962 Human bra
42	40	74.1	50 4 ABG53184	Abg53184 Human liv
43	40	74.1	50 5 ABG41302	Abg41302 Human pep
44	40	74.1	98 5 ABP07670	Abp07670 Human ORF
45	40	74.1	186 2 AAM69879	Aam69879 Antigen 2

## ALIGNMENTS

RESULT 1  
AAB33009  
ID AAB33009 standard; protein; 171 AA.  
XX  
AC AAB33009;  
DT 25-JAN-2001 (first entry)  
XX  
DE Pinus radiata transcription factor protein sequence #136.  
XX  
XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
XX poplar; sweetgum; teak; mahogany; bzlp; G-box binding factor;  
XX basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
XX homeodomain zipper; LIM domain; AP2; BREVES; zinc finger domain;  
XX type 2 Cys2His2; CCNAR box element; MYB.  
XX  
OS Pinus radiata.  
XX  
FN W0200053724-A2.  
XX  
PD 14-SEP-2000.  
XX  
PP 09-MAR-2000; 2000MO-US006112.  
XX  
PR 11-MAR-1999; 99US-00266513.  
XX 18-AUG-1999; 99US-0149485P.  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
XX  
PI Wood M, McGrath A, Shenk MA, Glenn M;  
XX WPI, 2000-579369/54.  
XX  
PS  
XX  
XX New isolated polynucleotide encoding a plant transcription factor for  
XX producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
XX having modified gene expression or modified activity of a polypeptide.  
XX  
XX Claim 8; Page 390; 747pp; English.  
XX  
XX The present invention relates to novel plant transcription factors from  
XX Eucalyptus grandis or Pinus radiata. The present sequence is one such  
XX transcription factor. The transcription factor may be used to produce a  
XX plant having modified gene expression such as a woody plant e.g. a  
XX eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or  
XX to modify the activity of a polypeptide in a plant. The transcription  
XX factors of the present invention are members from the following families  
XX of regulatory proteins: bzlp, bzlp family of G-box binding factors, basic

CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain  
 CC zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2  
 CC Cys2His2, CCAAT box elements and MYB  
 XX Sequence 171 AA;

Query Match 87.0%; Score 47; DB 3; Length 171;  
 Best Local Similarity 64.3%; Pred. No. 0.07;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGVITPYXXQVXXL 14  
 ||:|||||  
 Db 60 IGVITPYAAQVGL 73

## RESULT 2

ABW01183  
 ID ABW01183 standard; peptide; 14 AA.

XX ABW01183;

DT 15-JUN-2004 (first entry)

XX Saccharomyces cerevisiae motif VII peptide.

XX Modulator of translation termination; MTT1; helicase B; antiviral;  
 KM therapy; HCSB; nonsense mutation; yeast.

XX Saccharomyces cerevisiae.

OS Saccharomyces cerevisiae.

XX Key Location/Qualifiers

FT Misc-difference 8 /label= Unknown

FT Misc-difference 9 /note= "Xaa may be any amino acid"

FT Misc-difference 12 /note= "Xaa may be any amino acid"

FT Misc-difference 13 /note= "Xaa may be any amino acid"

FT Misc-difference 13 /label= Unknown

FT Misc-difference 13 /note= "Xaa may be any amino acid"

PN US630294-B1.

XX 07-OCT-2003.

XX 22-JUL-1999; 99US-00359268.

XX 22-JUL-1998; 98US-0093685P.

XX (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

XX Peltz S, Czaplinski K, Dimman JD;

XX WPI; 2003-810549/76.

PT Identifying an agent that increases nonsense suppression, for antiviral  
 PT therapy, by contacting modulator of translation termination (MTT1) in  
 PT Saccharomyces cerevisiae with a test agent, and detecting specific  
 PT binding to MTT1.

PS Disclosure; Col 43; opp; English.

XX The invention relates to a method of identifying an agent that increases  
 CC nonsense suppression, by contacting modulator of translation termination  
 CC (MTT1) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.

XX The method is useful for identifying compositions or agents which  
 CC increase nonsense suppression. The invention may also be used for  
 CC antiviral therapy and for suppression of pathological nonsense mutations.

XX The present sequence is Saccharomyces cerevisiae motif VII peptide

SQ Sequence 14 AA;

Query Match 85.2%; Score 46; DB 7; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 0.0064;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGVITPYXXQVXXL 14  
 ||:|||||  
 Db 1 IGVITPYXXQVXXL 14

## RESULT 3

ADP44109  
 ID ADP44109 standard; peptide; 14 AA.

XX ADP44109;

DT 18-NOV-2004 (first entry)

XX Yeast translation termination modulation protein motif VII.

XX gene therapy; translation termination; RNA helicase; MTT1;

XX frameshift frequency; aberrant transcript degradation;

XX peptidyl transferase modulation; beta-thalassemia; beta-globin;

XX Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;

XX Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;

XX Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic fibrosis;

XX Kidney Stone; Familial hypercholesterolaemia; Retinitis Pigmentosa;

XX Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; Yeast.

XX Saccharomyces cerevisiae.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 8.9 /note= "Any amino acid"

FT Misc-difference 12.13 /note= "Any amino acid"

FT Misc-difference 12.13 /note= "Any amino acid"

PN US2004115787-A1.

XX 17-JUN-2004.

XX 28-AUG-2003; 2003US-00652334.

XX 22-JUL-1998; 98US-0093685P.

XX 22-JUL-1999; 99US-00359268.

XX (PELTZ) PELTZ S.

XX (CZAP/) CZAPLINSKI K.

XX (DINM/) DINMAN J D.

XX Peltz S, Czaplinski K, Dimman JD;

XX WPI; 2004-449400/42.

PT Identifying a test composition or agent that modulates the efficiency of  
 PT translation termination comprises contacting the MTT1 with the test  
 PT composition or agent, and determining if the test composition or agent  
 PT inhibits the MTT1.

PS Claim 39; SEQ ID NO 7; 41bp; English.

XX The invention relates to a method of identifying a test composition that  
 CC modulates the efficiency of translation termination comprising contacting  
 CC the RNA helicase MTT1 with a composition or agent under conditions  
 CC permitting binding between the MTT1 and the composition, detecting  
 CC specific binding of the test composition or agent to the MTT1, and  
 CC determining if the test composition or agent inhibits the MTT1. The  
 CC composition and methods are useful for modulating the fidelity of  
 CC translation termination or for identifying agents that affect the  
 CC functional activity of mRNAs by altering frameshift frequency, permit  
 CC monitoring of a termination event, promote degradation of aberrant

CC	transcripts, and provide modulators (inhibitors/stimulators) of peptidyl transferase activity during initiation, elongation, termination and mRNA degradation of translation. The agents, which may be antagonists or agonists, are useful in screening, diagnostic and therapeutic purposes, for diseases or conditions resulting from or cause premature translation, such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease, Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour, Hirschsprung disease, Cystic fibrosis, Kidney Stones, Familial hypercholesterolemia, Retinitis pigmentosa, or Neurofibromatosis, Reninoplasmatoma, ATM or Costmann Disease. The present sequence represents the amino acid sequence of the yeast translation termination modulation protein motif VII.
CC	
CC	
CC	
XX	Sequence 14 AA;
SQ	
Query Match	85.2%; Score 46; DB 8; Length 14;
Best Local Similarity	100.0%; Pred. No. 0.0064;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
OY	1 IGVTPTXXXQVXXL 14       1 IGVITPXXXXVXXL 14
DB	
RESULT 4	
AAY77810	
ID	AAY77810 standard; peptide; 15 AA.
AC	
AA777810;	
XX	
DT	31-MAY-2000 (first entry)
XX	
DE	Motif VII comprised in a gene modulating translation termination.
XX	
KM	Hellcasee B; HGCB; MTI1; modulator of translation termination; ERP1; ERG3; eukaryotic release factor; peptidyl transferase; beta-thalassaemia; beta-globin; Duchene/Becker Muscular Dystrophy; anti-tamemic.
KW	
XX	
OS	Unidentified.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 1..15 /note= "residues indicated Xaa are unspecified"
PX	
PN	WO200005586-A2.
XX	
PD	03-FEB-2000.
XX	
PX	22-JUL-1999; 99WO-US016802.
XX	
PR	22-JUL-1998; 98US-00120435.
XX	
PA	(UTNE-) UNIV NEW JERSEY.
XX	
P1	Peltz S, Czaplinski K, Dimman JD,
DR	WPI; 2000-171458/15.
XX	
XX	New multiprotein complex which can modulate peptidyl transferase activity during translation, useful to treat diseases associated with peptidyl transferase activity e.g. Duchene/Becker Muscular Dystrophy.
PS	Claim 39; Page 80; 89pp; English.
CC	The invention provides a new multiprotein complex which can modulate peptidyl transferase activity during translation. The complex comprises the gene encoding Hellcasee B (HGCB, renamed MTI1, for Modulator of Translation Termination) and the conserved proteins known to interact and carry out translation termination in eukaryotic cells, peptidyl eukaryotic release factor (ERP) 1 and ERG3. The complex can be used to modulate peptidyl transferase activity during translation in a cell. It can be administered therapeutically combined with a carrier in

CC	pharmaceutical compositions to treat diseases associated with peptidyl
CC	transferase activity, especially diseases resulting from a nonsense or
CC	frameshift mutation e.g. beta-thalassemia, beta-globin, Duchene/Becker
CC	Muscular Dystrophy etc. It can be used to identify disease conditions
CC	Involving a defect in the complex, by transfecting cells with encoding
CC	nucleic acid and determining the proportion of defective complex before
CC	and after transfection. It is also useful to screen for drugs involved in
CC	peptidyl transferase activity during translation, inhibiting the
CC	interaction between mTOR and eRF3 or involved in enhancing translation
CC	termination. Vectors comprising polynucleotides encoding the complex (or
CC	antisense sequences) can be constructed and introduced into cells to
CC	interfere with complex expression and so modulate the efficiency of
CC	translation termination of mRNA and/or degradation of aberrant
CC	transcripts in a cell. Agents binding to the complex can be identified
CC	and included in therapeutic compositions useful as above, and/or used to
CC	modulate peptidyl transferase activity during translation in cells. They
CC	are also useful to modulate the efficiency of translation termination of
CC	mRNA at a nonsense codon and/or promote degradation of aberrant
CC	transcripts in cells. The method can be used to identify agents/
CC	compositions modulating binding to mTOR1, useful to identify genes.
CC	Sequences AA47804-912 represent motifs I-IX comprised in the genes of
CC	interest, used for modulating translation termination
XX	
SQ	Sequence 15 AA;
Query Match	85.2%; Score 46; DB 3; Length 15;
Best Local Similarity	100.0%; Pred. No. 0.0069;
Matches 14; Conservative	0; Mismatches 0; Indels 0; Gaps 0
Gy	1 IGVTTPYXXQVXXL 14       1 IGVTTPYXXQVXXL 14
Dd	-
RESULT 5	
ADS42927	
ID	ADS42927 standard; protein; 648 AA.
XX	
AC	ADS42927;
XX	
DT	02-DEC-2004 (first entry)
XX	
DE	Bacterial polypeptide #21357.
XX	
KW	Recombinant DNA construct; transformed plant; improved plant property;
KW	cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW	pathogen tolerance; pest tolerance; plant disease resistance;
KW	cell cycle pathway modification; plant growth regulator;
KW	homologous recombination; seed oil yield; protein yield; carbohydrate;
KW	nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX	bacterial polypeptide.
OS	Bacteria.
XX	
PN	US2003233675-A1.
XX	
PD	18-DEC-2003.
XX	
PP	20-FEB-2003; 2003US-00369493.
XX	
FR	21-FEB-2002; 2002US-0360039P.
XX	
PA	(CAOY/) CAO Y.
PA	(HINK/) HINKLE G J.
PA	(SLAT/) SLATER S C.
PA	(CHEN/) CHEN X.
PA	(GOLD/) GOLDMAN B S.
XX	
PI	Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX	
RR	WPI; 2004-061375/06.
XX	
PT	New recombinant DNA construct comprising a promoter positioned to provide

PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.  
 XX  
 PS Claim 1; SEQ ID NO 21357; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition, improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
 CC  
 XX  
 SQ Sequence 648 AA;

Query Match 85.2%; Score 46; DB 8; Length 648;  
 Best Local Similarity 71.4%; Pred. No. 0.54;  
 Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IGVITPYXXQVXXL 14  
 |||||  
 Db 537 IGVITPYDDQVDL 550

RESULT 6  
 ID ADN19868 standard; protein; 1944 AA.  
 XX  
 AC ADN19868;

XX 02-DEC-2004 (first entry)  
 XX

DE Bacterial polypeptide #2521.  
 XX

KW Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.  
 KM  
 XX  
 OS Bacteria.  
 XX  
 PN US2003233675-A1.  
 XX  
 PD 18-DEC-2003.  
 XX  
 PF 20-FEB-2003; 2003US-00369493.  
 XX  
 PR 21-FEB-2002; 2002US-036039P.  
 XX  
 XX (CAOY/) CAO Y.  
 PA (HINK/) HINKLE G J.  
 PA (SLAT/) SLATER S C.  
 PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.  
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
 PI  
 XX WPI; 2004-061375/06.  
 DR

XX New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.  
 XX  
 PS Claim 1; SEQ ID NO 2521; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition, improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
 CC  
 XX  
 SQ Sequence 1944 AA;

Query Match 85.2%; Score 46; DB 8; Length 1944;  
 Best Local Similarity 64.3%; Pred. No. 1.9;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IGVITPYXXQVXXL 14  
 |||||  
 Db 1671 IGVITPYRSQVQQL 1684

RESULT 7  
 ID ABB65781 standard; protein; 1274 AA.  
 XX  
 AC ABB65781;

XX 26-MAR-2002 (first entry)  
 XX

DE Drosophila melanogaster polypeptide SEQ ID NO 24135.  
 XX

KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 KM  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li FMD, Myers EW;  
 XX WPI; 2001-656860/75.  
 DR N-PSDB; ABL09884.  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 PS Disclosure; SEQ ID NO 24135; 21pp + Sequence listing; English.  
 XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signaling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-  
 CC AB572072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pcc\_sequences  
 XX  
 SQ Sequence 1274 AA;  
 Query Match 83.3%; Score 45; DB 4; Length 1274;  
 Best Local Similarity 64.3%; Pred. No. 2;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 Qy 1 IGVITPYXXQVXXL 14  
 Db 1131 IGVITPYXQVXXL 1144  
 RESULT 8  
 ADB63936  
 ID ADB63936 standard; protein; 125 AA.  
 XX  
 AC ADB63936;  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Human protein encoded by clone BRACE20200770.  
 XX  
 KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;  
 KW cell regeneration; membrane protein; signal transduction-related protein;  
 KW transcription-related protein; osteoporosis; neurological disease;  
 KW cancer; tumour.  
 XX  
 OS Homo sapiens.  
 PN EP1308459-A2.  
 XX  
 PD 07-MAY-2003.  
 XX  
 PF 28-MAR-2002; 2002EP-00007401.  
 XX  
 PR 05-NOV-2001; 2001JP-00379298.  
 PR 25-JAN-2002; 2002US-00350978.  
 XX  
 PA (HELI-) HELIX RES INST.  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX  
 DR WPI; 2003-450961/43.  
 DR N-PSDB; ADB61986.  
 XX  
 PT New polynucleotides and polypeptides, useful for developing a diagnostic  
 PT marker or medicines for regulation of their expression and activity, or  
 PT as targets of gene therapy.

XX Claim 1; Page; 222pp; English.  
 PS The invention discloses a polynucleotide comprising a sequence selected  
 CC from 1970 fully defined nucleotide sequences which encode novel  
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide  
 CC or its partial peptide, an antibody binding to the polypeptide or peptide  
 CC of the polynucleotide, immunologically assaying the polypeptide or  
 CC peptide of the polynucleotide by contacting the polypeptide or peptide  
 CC with the antibody of the encoded protein, and observing the binding  
 CC between the two, a transformant carrying the polynucleotide in an  
 CC expressible manner and an antisense polynucleotide. The oligonucleotide  
 CC is useful as a primer for synthesizing the polynucleotide, or as a probe  
 CC for detecting the polynucleotide. The polynucleotides and encoded  
 CC proteins are useful as pharmaceutical agents and many disease-related  
 CC genes may be included in them, for developing a diagnostic marker or  
 CC medicines for regulation of their expression and activity, or as targets  
 CC of gene therapy. The genes are involved in tissue and/or cell  
 CC regeneration. Membrane proteins, signal transduction-related proteins,  
 CC transcription-related proteins, disease-related proteins and genes  
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,  
 CC neurological diseases, cancer, tumours. The cDNA may be used to regulate  
 CC the activity or expression of the encoded protein to treat diseases. The  
 CC sequence presented is a protein of the invention. Note: Some of the  
 CC sequence data for this patent is not represented in the printed  
 CC specification, but is based on sequence information supplied by the  
 CC European Patent Office.  
 XX  
 SQ Sequence 125 AA;  
 Query Match 81.5%; Score 44; DB 7; Length 125;  
 Best Local Similarity 81.8%; Pred. No. 0.22;  
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 IGVITPYXXQV 11  
 Db 106 IGVITPYXQV 116  
 RESULT 9  
 AAY77816  
 ID AAY77816 standard; peptide; 414 AA.  
 XX  
 AC AAY77816;  
 XX  
 DT 31-MAY-2000 (first entry)  
 XX  
 DE Yeast D1p1 protein fragment.  
 XX  
 KW Helicase B; HCSB; MTT1; modulator of translation termination; eRF1; eRF3;  
 KW eukaryotic release factor; peptidyl transferase; beta-thalassemia;  
 KW beta-globin; Duchenne/Becker Muscular Dystrophy; anemic; yeast;  
 KW helicase; D1p1.  
 XX  
 OS Saccharomyces cerevisiae.  
 PN WO200005586-A2.  
 XX  
 PD 03-FEB-2000.  
 XX  
 PF 22-JUL-1999; 99WO-US016802.  
 XX  
 PR 22-JUL-1998; 98US-00120435.  
 XX  
 PA (UTNB-) UNIV NEW JERSEY.  
 XX  
 PI Peltz S, Czaplinski K, Dinman JD;  
 XX  
 DR WPI; 2000-171458/15.  
 DR  
 XX  
 PT New multiprotein complex which can modulate peptidyl transferase activity  
 PT during translation, useful to treat diseases associated with peptidyl  
 PT transferase activity e.g. Duchenne/Becker Muscular Dystrophy.

XX Example 1; Fig 1; 89pp; English.

PS The invention provides a new multiprotein complex which can modulate

XX peptidyl transferase activity during translation. The complex comprises

CC the gene encoding Helicase B (HCSB; renamed MT1), for Modulator of

CC Translation Termination) and the conserved proteins known to interact and

CC carry out translation termination in eukaryotic cells, peptidyl

CC eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to

CC modulate peptidyl transferase activity during translation in a cell. It

CC can be administered therapeutically combined with a carrier in

CC pharmaceutical compositions to treat diseases associated with peptidyl

CC transferase activity, especially diseases resulting from a nonsense or

CC frameshift mutation e.g. beta-thalassemia, beta-globin, Duchenne/Becker

CC Muscular Dystrophy etc. It can be used to identify disease conditions

CC involving a defect in the complex, by transfecting cells with encoding

CC nucleic acid and determining the proportion of defective complex before

CC and after transfection. It is also useful to screen for drugs involved in

CC peptidyl transferase activity during translation, inhibiting the

CC interaction between MT1 and eRF3 or involved in enhancing translation

CC termination. Vectors comprising polynucleotides encoding the complex (or

CC antisense sequences) can be constructed and introduced into cells to

CC interfere with complex expression and so modulate the efficiency of

CC translation termination of mRNA and/or degradation of aberrant

CC transcripts in a cell. Agents binding to the complex can be identified

CC and included in therapeutic compositions useful as above, and/or used to

CC modulate peptidyl transferase activity during translation in cells. They

CC are also useful to modulate the efficiency of translation termination or

CC mRNA at a nonsense codon and/or promote degradation of aberrant

CC transcripts in cells. The method can be used to identify agents/

CC compositions modulating binding to MT1, useful to identify genes.

CC Sequences AAY77813-817 represent protein fragments from yeast superfamily

CC group I helicases

XX

SQ Sequence 414 AA;

Query Match 81.5%; Score 44; DB 3; Length 414;

Best Local Similarity 64.3%; Pred. No. 0.88;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 IGVITPYXXQVXXL 14

Db 342 IGVISPYNAQVSHL 355

RESULT 10

ABW01204

ID ABW01204 standard; protein; 414 AA.

XX

AC ABW01204;

XX

DT 15-JAN-2004 (first entry)

XX

DE Saccharomyces cerevisiae D1P1 protein.

XX

KM Modulator of translation termination; MT1; helicase B; antiviral;

KW therapy; HCSB; nonsense mutation; yeast.

XX

OS Saccharomyces cerevisiae.

XX

PN US6630294-B1.

XX

PD 07-OCT-2003.

XX

PF 22-JUL-1999; 99US-00359268.

XX

PR 22-JUL-1998; 98US-0093685P.

XX

PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

XX

PI Peltz S, Czaplinski K, Dinman JD;

XX

DR WPI; 2003-810549/76.

XX Identifying an agent that increases nonsense suppression, for antiviral

PT therapy, by contacting modulator of translation termination (MT1) in

PT Saccharomyces cerevisiae with a test agent, and detecting specific

PT binding to MT1.

XX

PS Disclosure; Col 55-58; 0pp; English.

XX

CC The invention relates to a method of identifying an agent that increases

CC nonsense suppression, by contacting modulator of translation termination

CC (MT1) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.

CC The method is useful for identifying compositions or agents which

CC increase nonsense suppression. The invention may also be used for

CC antiviral therapy and for suppression of pathological nonsense mutations.

CC The present sequence is Saccharomyces cerevisiae D1P1 protein

XX

SQ Sequence 414 AA;

Query Match 81.5%; Score 44; DB 7; Length 414;

Best Local Similarity 64.3%; Pred. No. 0.88;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 IGVITPYXXQVXXL 14

Db 342 IGVISPYNAQVSHL 355

RESULT 11

ADP44130

ID ADP44130 standard; protein; 414 AA.

XX

AC ADP44130;

XX

DT 18-NOV-2004 (first entry)

XX

DE Yeast helicase D1P1.

XX

KM gene therapy; translation termination; RNA helicase; MT1;

KM frameshift frequency; aberrant transcript degradation;

KM peptidyl transferase modulation; beta-thalassemia; beta-globin;

KM Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;

KM Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;

KM Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic fibrosis;

KM Kidney Stone; Familial hypercholesterolaemia; Retinitis Pigmentosa;

KM Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; Yeast; enzyme.

XX

OS Saccharomyces cerevisiae.

XX

PN US2004115787-A1.

XX

PD 17-JUN-2004.

XX

PF 28-AUG-2003; 2003US-00652334.

XX

PR 22-JUL-1998; 98US-0093685P.

XX

PR 22-JUL-1999; 99US-00359268.

XX

PA (PELTZ) PELTZ S.

PA (CZAP/) CZAPLINSKI K.

PA (DINM/) DINMAN J D.

XX

PI Peltz S, Czaplinski K, Dinman JD;

XX

DR WPI; 2004-449400/42.

XX

PT Identifying a test composition or agent that modulates the efficiency of

PT translation termination comprises contacting the MT1 with the test

PT composition or agent, and determining if the test composition or agent

PT inhibits the MT1.

XX

PS Disclosure; SEQ ID NO 28; 41pp; English.

XX

DR The invention relates to a method of identifying a test composition that



CC modulates the efficiency of translation termination comprising contacting  
 CC the RNA helicase MTR1 with a composition or agent under conditions  
 CC permitting binding between the MTR1 and the composition, detecting  
 CC specific binding of the test composition or agent to the MTR1, and  
 CC determining if the test composition or agent inhibits the MTR1. The  
 CC composition and methods are useful for modulating the fidelity of  
 CC translation termination or for identifying agents that affect the  
 CC functional activity of mRNAs by altering frameshift frequency, permit  
 CC monitoring of a termination event, promote degradation of aberrant  
 CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl  
 CC transferase activity during initiation, elongation, termination and mRNA  
 CC degradation of translation. The agents, which may be antagonists or  
 CC agonists, are useful in screening, diagnostic and therapeutic purposes,  
 CC for diseases or conditions resulting from or cause premature translation,  
 CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular  
 CC Dystrophy, Haemophilia A, Haemophilia B, von Willebrand Disease,  
 CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,  
 CC Hirschsprung disease, Cystic fibrosis, Kidney Stones, Familial  
 CC hypercholesterolaemia, Retinitis Pigmentosa, or Neurofibromatosis,  
 CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents  
 CC the amino acid sequence of the yeast helicase Dpl1.

XX Sequence 414 AA;

Query Match 81.5%; Score 44; DB 8; Length 414;  
 Best Local Similarity 64.3%; Pred. No. 0.88;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGVITPYXXQVXXL 14  
 |||||  
 Db 342 IGVISPYNAQVSHL 355

RESULT 12

AB99784 ID ABB99784 standard; protein; 517 AA.

XX AC ABB99784;

XX DT 24-MAR-2003 (first entry)

XX DE Amino acid sequence of cardiac helicase activated by MEF2C protein.

XX KW CHAMP; cardiac helicase activated by MEF2 protein; cardiomyocyte;

XX KM cardiac function; heart; myocardial infarction; heart failure;

XX OS cardiac hypertrophy; exercise tolerance; cardiac hypertrophy.

XX OS Homo sapiens.

XX PN MO200295016-A2.

XX PD 28-NOV-2002.

XX PF 15-FEB-2002; 2002MO-US022511.

XX PR 16-FEB-2001; 2001US-0269764P.

XX PR 24-JAN-2002; 2002US-0351713P.

XX PA (TEXA ) UNIV TEXAS SYSTEM.

XX PI Olson E, Liu Z;

XX WP1; 2003-148465/14.

XX DR N-PSDB; ABZ23276.

XX PT New CHAMP (cardiac helicase activated by MEF2 protein) polypeptide and

XX PT polynucleotide, useful for enhancing cardiac function for treating,

XX PT inhibiting progression of or preventing myocardial infarction or heart

XX PT failure.

XX PS Claim 2; Page 132-133; 133pp; English.

XX CC The present sequence represents a polypeptide, designated CHAMP (cardiac

CC helicase activated by MEF2 protein). The CHAMP protein contains seven  
 CC conserved motifs bearing a striking resemblance to RNA helicases involved  
 CC in RNA processing, and to enhancer binding factors involved in tissue  
 CC transcription. CHAMP is expressed in cardiomyocytes from the linear tube  
 CC stage to adulthood. The CHAMP polynucleotide is useful for enhancing  
 CC cardiac function in a mammal. The CHAMP polypeptide and polynucleotide  
 CC are particularly useful for increasing CHAMP activity in the heart or  
 CC heart cells of a subject, in order to treat, inhibit progression of or  
 CC prevent myocardial infarction, heart failure or cardiac hypertrophy. This  
 CC is also useful for increasing exercise tolerance, reducing  
 CC hospitalization, improving quality of life, decreasing morbidity, or  
 CC decreasing mortality in a subject with heart failure or cardiac  
 CC hypertrophy

XX Sequence 517 AA;

Query Match 81.5%; Score 44; DB 6; Length 517;  
 Best Local Similarity 81.8%; Pred. No. 1.1;  
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGVITPYXXQV 11  
 |||||  
 Db 369 IGVITPYRKQV 379

RESULT 13

AB99781 ID ABB99781 standard; protein; 550 AA.

XX AC ABB99781;

XX DT 24-MAR-2003 (first entry)

XX DE Amino acid sequence of cardiac helicase activated by MEF2C protein.

XX KW CHAMP; cardiac helicase activated by MEF2 protein; cardiomyocyte;

XX KM cardiac function; heart; myocardial infarction; heart failure;

XX OS cardiac hypertrophy; exercise tolerance; cardiac hypertrophy.

XX OS Mus musculus.

XX PN MO200295016-A2.

XX PD 28-NOV-2002.

XX PF 15-FEB-2002; 2002MO-US022511.

XX PR 16-FEB-2001; 2001US-0269764P.

XX PR 24-JAN-2002; 2002US-0351713P.

XX PA (TEXA ) UNIV TEXAS SYSTEM.

XX PI Olson E, Liu Z;

XX WP1; 2003-148465/14.

XX DR N-PSDB; ABZ23273.

XX PT New CHAMP (cardiac helicase activated by MEF2 protein) polypeptide and

XX PT polynucleotide, useful for enhancing cardiac function for treating,

XX PT inhibiting progression of or preventing myocardial infarction or heart

XX PT failure.

XX PS Claim 2; Page 117-118; 133pp; English.

XX CC The present sequence represents a polypeptide, designated CHAMP (cardiac

XX CC helicase activated by MEF2 protein). The CHAMP protein contains seven

XX CC conserved motifs bearing a striking resemblance to RNA helicases involved

CC prevent myocardial infarction, heart failure or cardiac hypertrophy. This  
 CC is also useful for increasing exercise tolerance, reducing  
 CC hospitalization, improving quality of life, decreasing morbidity, or  
 CC decreasing mortality in a subject with heart failure or cardiac  
 CC hypertrophy

XX Sequence 550 AA;

Query Match 81.5%; Score 44; DB 6; Length 550;  
 Best Local Similarity 81.8%; Pred. No. 1.2;  
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGVITPYXXQV 11  
 |||||  
 Db 428 IGVITPYRKQV 438

RESULT 14  
 ADS43834  
 ID ADS43834 standard; protein; 683 AA.

XX ADS43834;

DT 02-DEC-2004 (first entry)

DE Bacterial polypeptide #22264.

XX Recombinant DNA construct; transformed plant; improved plant property;  
 KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;  
 KM pathogen tolerance; pest tolerance; plant disease resistance;  
 KM cell cycle pathway modification; plant growth regulator;  
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KM bacterial polypeptide.

XX Bacteria.

PN US2003233675-A1.

PD 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

PR 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 22264; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition, improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX Sequence 683 AA;

Query Match 81.5%; Score 44; DB 8; Length 683;  
 Best Local Similarity 64.3%; Pred. No. 1.6;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGVITPYXXQVXXL 14  
 |||||  
 Db 569 IGVISPYNAQVSHL 582

RESULT 15  
 ADB65007  
 ID ADB65007 standard; protein; 804 AA.

XX ADB65007;

DT 04-DEC-2003 (first entry)

DE Human protein encoded by clone PUEN20002470.

XX Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;  
 KM cell regeneration; membrane protein; signal transduction-related protein;  
 KM transcription-related protein; osteoporosis; neurological disease;  
 KM cancer; tumour.

OS Homo sapiens.

PN EPI308459-A2.

XX 07-MAY-2003.

XX 28-MAR-2002; 2002EP-00007401.

XX 05-NOV-2001; 2001JP-00379298.

XX 25-JAN-2002; 2002US-00350978.

XX (HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Tamechika I;

XX Seti N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;

XX WPI; 2003-450961/43.

XX N-PSDB; ADB63037.

XX New polynucleotides and polypeptides, useful for developing a diagnostic  
 PT marker or medicines for regulation of their expression and activity, or  
 PT as targets of gene therapy.

XX Claim 1; Page; 222pp; English.

XX The invention discloses a polynucleotide comprising a sequence selected  
 CC from 1970 fully defined nucleotide sequences which encode novel  
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide  
 CC or its partial peptide, an antibody binding to the polypeptide or peptide  
 CC of the polynucleotide, immunologically assaying the polypeptide or  
 CC peptide of the polynucleotide by contacting the polypeptide or peptide  
 CC with the antibody of the encoded protein, and observing the binding

CC between the two, a transformant carrying the polynucleotide in an  
CC expressible manner and an antisense polynucleotide. The oligonucleotide  
CC is useful as a primer for synthesizing the polynucleotide, or as a probe  
CC for detecting the polynucleotide. The polynucleotides and encoded  
CC proteins are useful as pharmaceutical agents and many disease-related  
CC genes may be included in them, for developing a diagnostic marker or  
CC medicines for regulation of their expression and activity, or as targets  
CC of gene therapy. The genes are involved in tissue and/or cell  
CC regeneration. Membrane proteins, signal transduction-related proteins,  
CC transcription-related proteins, disease-related proteins and genes  
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,  
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate  
CC the activity or expression of the encoded protein to treat diseases. The  
CC sequence presented is a protein of the invention. Note: Some of the  
CC sequence data for this patent is not represented in the printed  
CC specification, but is based on sequence information supplied by the  
CC European Patent Office.

XX  
SQ Sequence 804 AA;

Query Match 81.5%; Score 44; DB 7; Length 804;  
Best Local Similarity 81.8%; Pred. No. 1.9;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IGVITPYXXOV 11  
|||  
|||  
Db 702 IGVITPYRKQV 712

Search completed: April 18, 2005, 08:03:50  
Job time : 71.8045 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 07:37:21 ; Search time 17.4745 Seconds  
(without alignments)  
59.807 Million cell updates/sec

Title: US-10-652-334-7  
Perfect score: 54  
Sequence: 1 IGVITPYXXQVXXL 14

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	87.0	171	US-09-640-211A-1058	Sequence 1058, Ap
2	46	85.2	14	US-09-359-268A-7	Sequence 7, Appli
3	44	81.5	414	US-09-359-268A-28	Sequence 28, Appli
4	44	81.5	683	US-09-538-092-483	Sequence 483, App
5	43	79.6	426	US-09-248-796A-15170	Sequence 15170, A
6	42	77.8	415	US-09-359-268A-25	Sequence 25, Appli
7	42	77.8	1176	US-09-489-039A-8879	Sequence 8879, Ap
8	41	75.9	10	US-09-359-268A-11	Sequence 11, Appli
9	41	75.9	238	US-09-270-767-56745	Sequence 56745, A
10	41	75.9	486	US-09-270-767-41521	Sequence 41521, A
11	40	74.1	380	US-09-358-268A-29	Sequence 29, Appli
12	40	74.1	917	US-09-248-796A-19347	Sequence 19347, A
13	40	74.1	971	US-08-724-354D-22	Sequence 22, Appli
14	40	74.1	971	US-09-270-984A-22	Sequence 22, Appli
15	40	74.1	971	US-09-177-431-8	Sequence 8, Appli
16	40	74.1	3177	US-08-477-451-4	Sequence 4, Appli
17	39	72.2	1043	US-08-724-354D-4	Sequence 4, Appli
18	39	72.2	1043	US-09-270-984A-4	Sequence 4, Appli
19	39	72.2	1118	US-08-724-354D-2	Sequence 2, Appli
20	39	72.2	1118	US-09-270-984A-2	Sequence 2, Appli
21	39	72.2	1140	US-09-949-016-10116	Sequence 10116, A
22	38	70.4	405	US-09-248-796A-19107	Sequence 19107, A
23	36	66.7	211	US-09-902-540-16148	Sequence 16148, A
24	36	66.7	472	US-09-359-268A-26	Sequence 26, Appli
25	35	64.8	85	US-09-270-767-59862	Sequence 59862, A
26	35	64.8	444	US-09-270-767-44429	Sequence 44429, A
27	35	64.8	993	US-09-538-092-1100	Sequence 1100, Ap

28	34	63.0	200	4	US-09-902-540-13130	Sequence 13130, A
29	34	63.0	241	4	US-09-581-345-5	Sequence 5, Appli
30	33	61.1	6	4	US-09-359-268A-20	Sequence 20, Appli
31	33	61.1	153	4	US-08-311-731A-265	Sequence 265, App
32	33	61.1	181	4	US-09-248-796A-16115	Sequence 16115, A
33	33	61.1	231	4	US-09-198-452A-419	Sequence 419, App
34	33	61.1	260	4	US-09-438-185A-402	Sequence 402, App
35	33	61.1	547	4	US-09-489-039A-13843	Sequence 13843, A
36	33	61.1	774	4	US-09-252-991A-24623	Sequence 24623, A
37	32	59.3	152	4	US-09-107-433-1408	Sequence 1408, Ap
38	32	59.3	156	4	US-09-489-039A-12183	Sequence 12183, A
39	32	59.3	219	4	US-09-270-767-44821	Sequence 44821, A
40	32	59.3	279	4	US-09-134-000C-5430	Sequence 5430, Ap
41	32	59.3	344	4	US-09-489-039A-11458	Sequence 11458, A
42	32	59.3	1033	4	US-09-583-110-5097	Sequence 5097, Ap
43	32	59.3	1042	3	US-09-387-695-2	Sequence 2, Appli
44	32	59.3	2396	1	US-08-157-005-2	Sequence 2, Appli
45	32	59.3	2396	1	US-08-747-863-2	Sequence 2, Appli

## ALIGNMENTS

```
RESULT 1
US-09-640-211A-1058
; Sequence 1058, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640.211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ. ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1058
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-640-211A-1058

Query Match      87.0%; Score 47; DB 4; Length 171;
Best Local Similarity 64.3%; Pred. No. 0.045;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Cy      1 IGVITPYXXQVXXL 14
Db      60 IGVITPYAAQVGLL 73

RESULT 2
US-09-359-268A-7
; Sequence 7, Application US/09359268A
; Patent No. 6630294
; GENERAL INFORMATION:
; APPLICANT: Peltz, Stuart
; APPLICANT: Czaplinski, Kevin
; APPLICANT: Dinman, Jonathan D.
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
; FILE REFERENCE: 601-1-85N
; CURRENT APPLICATION NUMBER: US/09/359.268A
; CURRENT FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,685
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ. ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
```



Db 306 IGVITPYSAQ 315

## RESULT 7

US-09-489-039A-8879  
; Sequence 8879, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489, 039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 8879  
; LENGTH: 1176  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-8879

Query Match 77.8%; Score 42; DB 4; Length 1176;  
Best Local Similarity 42.9%; Pred. No. 3.9;  
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGVITPYXXOVXL 14

Db 888 VGVVTFPSAQVSTI 901

## RESULT 8

US-09-359-268A-11  
; Sequence 11, Application US/09359268A  
; Patent No. 6630294  
; GENERAL INFORMATION:  
; APPLICANT: Peltz, Stuart  
; APPLICANT: Czaplinski, Kevin  
; APPLICANT: Dielman, Jonathan D.  
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
; FILE REFERENCE: 601-1-85N  
; CURRENT APPLICATION NUMBER: US/09/359, 268A  
; CURRENT FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 60/093, 685  
; PRIOR FILING DATE: 1998-07-22  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: saccharomyces cerevisiae  
; FEATURE:  
; OTHER INFORMATION: Xaa = any amino acid  
US-09-359-268A-11

Query Match 75.9%; Score 41; DB 4; Length 10;  
Best Local Similarity 90.0%; Pred. No. 0.035;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IGVITPYXXQ 10

Db 1 IGVITPYXAQ 10

## RESULT 9

US-09-270-767-56745  
; Sequence 56745, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270, 767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 56745  
; LENGTH: 238  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-56745

Query Match 75.9%; Score 41; DB 4; Length 238;  
Best Local Similarity 63.6%; Pred. No. 1.1;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGVITPYXXOV 11

Db 125 IGVITPYSDQV 135

## RESULT 10

US-09-270-767-41521  
; Sequence 41521, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270, 767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 41521  
; LENGTH: 486  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-41521

Query Match 75.9%; Score 41; DB 4; Length 486;  
Best Local Similarity 63.6%; Pred. No. 2.4;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGVITPYXXOV 11

Db 126 IGVITPYSDQV 136

## RESULT 11

US-09-359-268A-29  
; Sequence 29, Application US/09359268A  
; Patent No. 6630294  
; GENERAL INFORMATION:  
; APPLICANT: Peltz, Stuart  
; APPLICANT: Czaplinski, Kevin  
; APPLICANT: Dielman, Jonathan D.  
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
; FILE REFERENCE: 601-1-85N  
; CURRENT APPLICATION NUMBER: US/09/359, 268A  
; CURRENT FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 60/093, 685  
; PRIOR FILING DATE: 1998-07-22  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 29  
; LENGTH: 380  
; TYPE: PRT  
; ORGANISM: saccharomyces cerevisiae  
US-09-359-268A-29

Query Match 74.1%; Score 40; DB 4; Length 380;  
Best Local Similarity 80.0%; Pred. No. 2.9;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGVITPYXXQ 10  
|||  
303 IGVITPYEQ 312

RESULT 12  
US-09-248-796A-19347  
Sequence 19347, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
PRIOR FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 19347  
LENGTH: 917  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-19347

Query Match 74.1%; Score 40; DB 4; Length 917;  
Best Local Similarity 70.0%; Pred. No. 7.6;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGVITPYXXQ 10  
|||  
777 IGVITPYEQ 786

RESULT 13  
US-08-724-354D-22  
Sequence 22, Application US/08724354D  
Patent No. 5994119  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,354D  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/016,482  
FILING DATE: 29-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 971 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-724-354D-22

Query Match 74.1%; Score 40; DB 2; Length 971;  
Best Local Similarity 80.0%; Pred. No. 8.1;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGVITPYXXQ 10  
|||  
726 IGVITPYEQ 735

RESULT 14  
US-09-270-984A-22  
Sequence 22, Application US/09270984A  
Patent No. 6048965  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/270,984A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/724,354  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 971 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-09-270-984A-22

Query Match 74.1%; Score 40; DB 3; Length 971;  
Best Local Similarity 80.0%; Pred. No. 8.1;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGVITPYXXQ 10  
|||  
726 IGVITPYEQ 735

RESULT 15  
US-09-177-431-8  
Sequence 8, Application US/09177431



; Patent No. 6071700  
; GENERAL INFORMATION:  
; APPLICANT: He, Feng  
; APPLICANT: Jacobson, Allan S.  
; TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE  
; TITLE OF INVENTION: ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/177,431  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/955,472  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Passer, J. Peter  
; REGISTRATION NUMBER: 32,983  
; REFERENCE/DOCKET NUMBER: 07917/050001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-9806  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 971 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: Internal  
; US-09-177-431-8

Query Match 74.1%; Score 40; DB 3; Length 971;  
Best Local Similarity 80.0%; Pred. No. 8.1;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGVITPYXQ 10  
|||  
Db 726 IGVITPYEQ 735

Search completed: April 18, 2005, 08:18:36  
Job time: 18.4745 secs

This page blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 08:06:16 ; Search time 49.6642 Seconds  
(Without alignments)  
93.693 Million cell updates/sec

Title: US-10-652-334-7  
Perfect score: 54  
Sequence: 1 IGVITPYXXQVXXL 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubppaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubppaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	47	87.0	127	US-10-767-701-48804
2	47	87.0	332	US-10-425-114-38637
3	47	87.0	404	US-10-425-114-37833
4	47	87.0	626	US-10-437-963-166322
5	47	87.0	637	US-10-424-599-233501
6	47	87.0	639	US-10-425-114-37127
7	47	87.0	642	US-10-425-114-37557
8	46	85.2	14	US-10-652-334-7
9	46	85.2	648	US-10-369-493-21357
10	46	85.2	1944	US-10-369-493-2521
11	45	83.3	182	US-10-767-701-62715
12	45	83.3	317	US-10-425-114-42716
13	44	81.5	69	US-10-437-963-106566

14	44	81.5	125	US-10-104-047-2090	Sequence 2090, Ap
15	44	81.5	309	US-10-425-114-39052	Sequence 39052, A
16	44	81.5	414	US-10-652-334-28	Sequence 28, Ap1
17	44	81.5	417	US-10-116-275-139	Sequence 139, App
18	44	81.5	517	US-10-077-583-8	Sequence 8, Ap11
19	44	81.5	550	US-10-077-583-2	Sequence 2, Ap11
20	44	81.5	683	US-10-369-493-22264	Sequence 22264, A
21	44	81.5	804	US-10-104-047-3161	Sequence 3161, Ap
22	44	81.5	959	US-10-437-963-123484	Sequence 123484, Ap
23	44	81.5	1187	US-09-801-574-14	Sequence 14, Ap1
24	44	81.5	1208	US-10-077-583-4	Sequence 4, Ap11
25	44	81.5	1211	US-09-801-574-60	Sequence 60, Ap11
26	44	81.5	1349	US-10-077-583-6	Sequence 6, Ap11
27	44	81.5	1975	US-10-437-963-140079	Sequence 140079, A
28	43	79.6	250	US-10-424-599-259854	Sequence 259854, A
29	43	79.6	642	US-10-369-493-21526	Sequence 21526, A
30	43	79.6	650	US-10-369-493-2884	Sequence 2884, Ap
31	43	79.6	1361	US-10-437-963-165703	Sequence 165703, A
32	43	79.6	2646	US-10-437-963-189923	Sequence 189923, A
33	42	77.8	415	US-10-652-334-25	Sequence 25, Ap1
34	41	75.9	10	US-10-652-334-11	Sequence 11, Ap1
35	41	75.9	219	US-10-767-701-57601	Sequence 57601, A
36	41	75.9	237	US-10-424-599-202249	Sequence 202249, A
37	40	74.1	50	US-09-864-761-38825	Sequence 38825, A
38	40	74.1	142	US-10-335-977-8637	Sequence 8637, Ap
39	40	74.1	172	US-10-424-599-278250	Sequence 278250, A
40	40	74.1	199	US-10-767-701-38822	Sequence 38822, A
41	40	74.1	215	US-10-425-114-42518	Sequence 42518, A
42	40	74.1	312	US-10-425-114-45797	Sequence 45797, A
43	40	74.1	350	US-10-424-599-241211	Sequence 241211, A
44	40	74.1	380	US-10-652-334-29	Sequence 29, Ap1
45	40	74.1	638	US-10-437-963-181690	Sequence 181690, A

#### ALIGNMENTS

RESULT 1  
US-10-767-701-48804  
Sequence 48804, Application US/10767701  
Publication No. US20040172684A1  
GENERAL INFORMATION:  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53535)B  
CURRENT APPLICATION NUMBER: US/10/767,701  
CURRENT FILING DATE: 2004-01-29  
NUMBER OF SEQ ID NOS: 63128  
SEQ ID NO 48804  
LENGTH: 127  
TYPE: PRT  
ORGANISM: Sorghum bicolor  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(127)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3476-013-Pl-K1-B11.pep  
US-10-767-701-48804  
Query Match 87.0%; Score 47; DB 16; Length 127;  
Best Local Similarity 64.3%; Pred. No. 0.15;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
Cy 1 IGVITPYXXQVXXL 14  
||:|||||  
Db 15 IGVITPYSAQVTC 28

RESULT 2

US-10-425-114-38637  
; Sequence 38637, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ. ID NOS: 73128  
; SEQ ID NO 38637  
; LENGTH: 332  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700241578\_FLI.pep  
US-10-425-114-38637

Query Match 87.0%; Score 47; DB 15; Length 332;  
Best Local Similarity 64.3%; Pred. No. 0.41;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGVITPYXXQVXXL 14  
||:|||||  
Db 220 IGIITPYAAQVTCL 233

RESULT 3  
US-10-425-114-37833  
; Sequence 37833, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ. ID NOS: 73128  
; SEQ ID NO 37833  
; LENGTH: 404  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700618821\_FLI.pep  
US-10-425-114-37833

Query Match 87.0%; Score 47; DB 15; Length 404;  
Best Local Similarity 64.3%; Pred. No. 0.51;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGVITPYXXQVXXL 14  
||:|||||  
Db 292 IGIITPYAAQVTCL 305

RESULT 4  
US-10-437-963-166322  
; Sequence 166322, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ. ID NOS: 204966  
; SEQ ID NO 166322  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(626)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_65042C.1.pep  
US-10-437-963-166322

Query Match 87.0%; Score 47; DB 16; Length 626;  
Best Local Similarity 64.3%; Pred. No. 0.81;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGVITPYXXQVXXL 14  
||:|||||  
Db 515 IGIITPYAAQVTCL 528

RESULT 5  
US-10-424-599-233501  
; Sequence 233501, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ. ID NOS: 285684  
; SEQ ID NO 233501  
; LENGTH: 637  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(637)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_52878C.1.pep  
US-10-424-599-233501

Query Match 87.0%; Score 47; DB 15; Length 637;  
Best Local Similarity 64.3%; Pred. No. 0.82;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGVITPYXXQVXXL 14  
||:|||||  
Db 525 IGIITPYAAQVTCL 538

RESULT 6  
US-10-425-114-37717  
; Sequence 37717, Application US/10425114

```
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovacic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 37717
/ LENGTH: 639
/ TYPE: PRT
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3170-056-H12_FLI.pep
US-10-425-114-37717

Query Match      87.0%; Score 47; DB 15; Length 639;
Best Local Similarity 64.3%; Pred. No. 0.83;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Cy 1 IGVITPYXXQVXXL 14
   |||||
Db 527 IGVITPYAAQVVL 540

RESULT 7
US-10-425-114-37557
/ Sequence 37557, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovacic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 37557
/ LENGTH: 642
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana columbia
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3175-019-D10_FLI.pep
US-10-425-114-37557

Query Match      87.0%; Score 47; DB 15; Length 642;
Best Local Similarity 64.3%; Pred. No. 0.83;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Cy 1 IGVITPYXXQVXXL 14
   |||||
Db 531 IGVITPYAAQVVL 544

RESULT 8
US-10-652-334-7
/ Sequence 7, Application US/10652334
/ Publication No. US2004011587A1
/ GENERAL INFORMATION:
/ APPLICANT: Pelcz, Stuart
/ APPLICANT: Czaplinski, Kevin
/ APPLICANT: Dimman, Jonathan D.
```

```
/ TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
/ TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
/ TITLE OF INVENTION: THEREOF
/ FILE REFERENCE: 601-1-85N
/ CURRENT APPLICATION NUMBER: US/10/652,334
/ CURRENT FILING DATE: 2003-08-28
/ PRIOR APPLICATION NUMBER: US/09/359,268A
/ PRIOR FILING DATE: 1999-07-22
/ PRIOR APPLICATION NUMBER: 60/093,685
/ PRIOR FILING DATE: 1998-07-22
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 7
/ LENGTH: 14
/ TYPE: PRT
/ ORGANISM: Saccharomyces cerevisiae
/ FEATURE:
/ OTHER INFORMATION: Xaa = any amino acid
US-10-652-334-7

Query Match      85.2%; Score 46; DB 16; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 IGVITPYXXQVXXL 14
   |||||
Db 1 IGVITPYXXQVXXL 14

RESULT 9
US-10-369-493-21357
/ Sequence 21357, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 21357
/ LENGTH: 648
/ TYPE: PRT
/ ORGANISM: Archaeoglobus fulgidus
US-10-369-493-21357

Query Match      85.2%; Score 46; DB 15; Length 648;
Best Local Similarity 71.4%; Pred. No. 1.3;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Cy 1 IGVITPYXXQVXXL 14
   |||||
Db 537 IGVITPYDDQVVL 550

RESULT 10
US-10-369-493-2521
/ Sequence 2521, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
```

```
;; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
;; FILE REFERENCE: 38-10(52052)B
;; CURRENT APPLICATION NUMBER: US/10/369,493
;; CURRENT FILING DATE: 2003-02-28
;; PRIOR APPLICATION NUMBER: US 60/360,039
;; PRIOR FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 2521
;; LENGTH: 1944
;; TYPE: PRT
;; ORGANISM: Schizosaccharomyces pombe
US-10-369-493--2521

Query Match
Best Local Similarity 85.2%; Score 46; DB 15; Length 1944;
Best Local Similarity 64.3%; Pred. No. 4.3;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Cy 1 IGVTPYXXQVXXL 14
Db 1671 IGVTPYRSQVQQL 1684

RESULT 11
US-10-767-701-62775
; Sequence 62775, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 62775
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 18065480.pcp
US-10-767-701-62775

Query Match
Best Local Similarity 83.3%; Score 45; DB 16; Length 182;
Best Local Similarity 64.3%; Pred. No. 0.55;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Cy 1 IGVTPYXXQVXXL 14
Db 27 IGVTPYRSQVQQL 40

RESULT 12
US-10-425-114-42716
; Sequence 42716, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven B
; APPLICANT: Tabaska, Jack B
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53131)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 42716
; LENGTH: 317
; TYPE: PRT
```

```
;; ORGANISM: Zea mays
;; FEATURE:
;; OTHER INFORMATION: Clone ID: 700469944_FLI.pcp
US-10-425-114-42716

Query Match
Best Local Similarity 83.3%; Score 45; DB 15; Length 317;
Best Local Similarity 64.3%; Pred. No. 0.99;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Cy 1 IGVTPYXXQVXXL 14
Db 49 IGVTPYRQVAKI 62

RESULT 13
US-10-437-963-106566
; Sequence 106566, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 106566
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_10999C.1.pcp
US-10-437-963-106566

Query Match
Best Local Similarity 81.5%; Score 44; DB 16; Length 69;
Best Local Similarity 72.7%; Pred. No. 0.31;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Cy 1 IGVTPYXXQV 11
Db 13 IGITTPYNAQV 23

RESULT 14
US-10-104-047-2090
; Sequence 2090, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1 full length cDNA
; FILE REFERENCE: HL-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2090
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2090

Query Match
Best Local Similarity 81.5%; Score 44; DB 15; Length 125;
Best Local Similarity 81.8%; Pred. No. 0.58;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

OY 1 IGVITPYXXQV 11  
|||  
Db 106 IGVITPYRKQV 116

RESULT 15  
US-10-425-114-39052  
; Sequence 39052, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 39052  
; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3180-045-F7\_FLI.pep  
US-10-425-114-39052

Query March 81.5%; Score 44; DB 15; Length 309;  
Best Local Similarity 72.7%; Pred. No. 1.5;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
OY 1 IGVITPYXXQV 11  
|||  
Db 194 IGVITPYNAQV 204

Search completed: April 18, 2005, 09:04:07  
Job time : 49.6642 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: Apr11 18, 2005, 07:32:05 ; Search time 13.4891 Seconds  
(without alignments)  
99.861 Million cell updates/sec

Title: US-10-652-334-7  
Perfect score: 54  
Sequence: 1 IGVITPYXXQVXXL 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	87.0	635	2 T02699	probable helicase
2	46	85.2	648	2 C69423	DNA helicase homol
3	46	85.2	1944	2 T40065	RNA-splicing endo
4	44	81.5	678	2 T42668	hypothetical prote
5	44	81.5	683	2 S34700	probable purine nu
6	44	81.5	1048	2 C66189	protein T25N20.11
7	44	81.5	1687	2 T38072	DNA2-NM7 helicase
8	43	79.6	245	2 D72258	helicase-related p
9	43	79.6	642	2 D69085	transcription cont
10	43	79.6	650	2 G72429	hypothetical prote
11	43	79.6	1825	2 T52521	related to SEN1 pr
12	43	79.6	2142	2 D66303	hypothetical prote
13	42	77.8	1339	2 D90303	hypothetical prote
14	42	77.8	530	2 D70476	DNA helicase - Agu
15	42	77.8	660	2 F85069	hypothetical prote
16	42	77.8	1121	2 S30862	DNA dependent ATPa
17	41	75.9	338	2 S56533	hypothetical 38K p
18	41	75.9	814	2 T00740	hypothetical prote
19	41	75.9	1075	2 C66682	protein F1B22.16 l
20	41	75.9	1076	2 B96682	protein F1B22.14 l
21	41	75.9	1090	2 T00533	probable DNA2-NM7
22	41	75.9	1311	2 T08986	hypothetical prote
23	40	74.1	274	2 D64588	probable DNA helic
24	40	74.1	653	2 B75105	probable DNA-bind
25	40	74.1	656	2 E71080	probable DNA-bind
26	40	74.1	971	2 S23408	prematurely termin
27	39	72.2	256	2 D66874	ABC transporter AT
28	39	72.2	453	2 G69494	DNA helicase homol
29	39	72.2	660	2 T41580	probable dna-bind

30	39	72.2	1004	2 A39611	probable GTP-bind
31	38	70.4	935	2 S62476	hypothetical prote
32	37	68.5	663	2 H64312	probable DNA helic
33	37	68.5	821	2 C64304	DNA helicase [limp
34	36	66.7	591	2 S64727	protein secretion
35	36	66.7	693	2 T26415	hypothetical prote
36	36	66.7	1118	2 S75309	hypothetical prote
37	36	66.7	2231	2 S53416	SEN1 protein - Yea
38	35	64.8	206	2 AG2521	hypothetical prote
39	35	64.8	952	2 AP2058	ABC transporter AT
40	35	64.8	989	2 T48845	Insulin II gene en
41	35	64.8	993	2 A47500	Ig mu chain switch
42	35	64.8	993	2 S35633	DNA-binding protei
43	35	64.8	1077	2 T50697	hypothetical prote
44	34	63.0	72	2 C37263	Ig heavy chain V r
45	34	63.0	102	2 B37263	Ig heavy chain V r

## ALIGNMENTS

RESULT 1  
T02699  
probable helicase At2g03270 [imported] - Arabidopsis thaliana  
N:Alternate names: hypothetical protein T1B12.6  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C:Accession: T02699; D84446  
R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shee, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Rom  
submitted to the EMBL Data Library, September 1998  
A:Description: Arabidopsis thaliana chromosome II BAC T1B12 genomic sequence.  
A:Reference number: Z14702  
A:Accession: T02699  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-635 <ROU>  
A:Cross-references: UNIPROT:O81047; EMBL:AC005313; NID:g3548797; PID:g3548803  
A:Experimental source: cultivar Columbia  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Bonito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Unayam, L.; Tallon, L.;  
euser, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: D84446  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-635 <STO>  
A:Cross-references: GB:AB002093; NID:g4335770; PID:AA017447.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g03270; T1B12.6  
A:Map position: 2  
C:Superfamily: probable DNA helicase MJ0104  
Query Match 87.0%; Score 47; DB 2; Length 635;  
Best local similarity 64.3%; Pred. No. 0.11;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 IGVITPYXXQVXXL 14  
Db 524 IGVITPYXXQVXXL 537  
RESULT 2  
C69423  
DNA helicase homolog - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: C69423  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997

A:Authors: Uteerback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Weese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea  
A:Reference number: A69250; MUID:98043343; PMID:9389475  
A:Accession: C69423  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-648 <KIE>  
A:Cross-references: UNIPROT:O28883; GB:AE001009; GB:AE000782; NID:g2689332; PIDN:AA8986  
C:Superfamily: probable DNA helicase MJ0104

Query Match 85.2%; Score 46; DB 2; Length 648;  
Best Local Similarity 71.4%; Pred. No. 0.17;  
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 IGVTPYXXQVXXL 14  
Db 537 IGVTPYDDQVLL 550

## RESULT 3

T40065  
tRNA-splicing endonuclease positive effector - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T40065  
R:Lyme, M.; Rajandream, M.A.; Bartell, B.G.; Lelaure, V.; Galibert, F.  
submitted to the EMBL Data Library, December 1998

A:Reference number: Z21903  
A:Accession: T40065  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1944 <LYN>  
A:Cross-references: UNIPROT:O94387; EMBL:AL034463; PIDN:CAA22438.1; GSPDB:GN00067; SPDB:  
A:Experimental source: strain 972h-; cosmid c29A10  
C:Genetics:  
A:Gene: SPDB:SPBC29A10.10c  
A:Map position: 2

Query Match 85.2%; Score 46; DB 2; Length 1944;  
Best Local Similarity 64.3%; Pred. No. 0.57;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 IGVTPYXXQVXXL 14  
Db 1671 IGVTPYRSQVQL 1684

## RESULT 4

T42668  
hypotheoretical protein DKFZp434B0717.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T42668  
R:Bioecker, H.; Bioecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, November 1999  
A:Reference number: Z22230  
A:Accession: T42668  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-678 <AAA>  
A:Cross-references: UNIPROT:Q9BXT6; EMBL:AL133068  
A:Experimental source: adult testis; clone DKFZp434B0717  
C:Genetics:  
A:Note: DKFZp434B0717.1

Query Match 81.5%; Score 44; DB 2; Length 678;  
Best Local Similarity 81.8%; Pred. No. 0.48;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 IGVTPYXXQV 11  
Db 530 IGVTPYRKQV 540

RESULT 5  
S34700  
probable purine nucleotide-binding protein YKL017c - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: S34700; S37830; S37834  
R:Wiemann, S.; Voss, H.; Schwager, C.; Rupp, T.; Stegmann, J.; Zimmermann, J.; Grothues,  
submitted to the EMBL Data Library, July 1993  
A:Description: Sequencing and analysis of 51.5 kilobases on the left arm of chromosome XI  
A:Reference number: S34679  
A:Accession: S34700  
A:Molecule type: DNA  
A:Residues: 1-683 <WIE>  
A:Cross-references: UNIPROT:P34243; EMBL:X74152; NID:g450363; PID:g395256  
A:Experimental source: strain S288C  
R:Wiemann, S.; Voss, H.; Schwager, C.; Rupp, T.; Grothues, D.; Seneen, C.; Stegmann, J.,  
submitted to the Protein Sequence Database, March 1994  
A:Reference number: S37825  
A:Accession: S37830  
A:Molecule type: DNA  
A:Residues: 1-683 <W12>  
A:Cross-references: EMBL:Z28017; NID:g486006; PID:g486007; MIPS:YKL017C  
A:Experimental source: strain S288C  
R:Rieger, M.  
submitted to the Protein Sequence Database, March 1994  
A:Reference number: S37832  
A:Accession: S37834  
A:Molecule type: DNA  
A:Residues: 1-683 <R1E>  
A:Cross-references: EMBL:Z28017; NID:g486006; PID:g486007; MIPS:YKL017C  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:DIP1  
A:Cross-references: SGD:S0001500; MIPS:YKL017C  
A:Map position: 11L  
C:Superfamily: probable DNA helicase MJ0104  
C:Keywords: ATP; P-loop; purine nucleotide binding  
F:229-236/Region: nucleotide-binding motif A (P-loop)  
F:235/Binding site: ATP/GTP (lys) #status predicted

Query Match 81.5%; Score 44; DB 2; Length 683;  
Best Local Similarity 64.3%; Pred. No. 0.48;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 IGVTPYXXQVXXL 14  
Db 569 IGVISPYNAQVSHL 582

## RESULT 6

C86189  
protein T25N20.11 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: C86189  
R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marzalli,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, I.  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.; Marzalli,  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: C86189  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1048 <STO>  
A:Cross-references: UNIPROT:Q9JLR8; GB:AE005172; NID:g8778728; PIDN:AAF79736.1; GSPDB:GN0

Query Match 81.5%; Score 44; DB 2; Length 683;  
Best Local Similarity 64.3%; Pred. No. 0.48;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 IGVTPYXXQVXXL 14  
Db 569 IGVISPYNAQVSHL 582

C:Genetics:  
A:Gene: T25N20.11  
A:Map position: 1

Query Match 81.5%; Score 44; DB 2; Length 1048;  
Best Local Similarity 81.8%; Pred. No. 0.77;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGVITPYXXQV 11  
|||  
Db 755 IGVITPYRQGV 765

## RESULT 7

T39072  
DNA2-NAN7 helicase family protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: T39072

R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Connor, R.E.; Wood, V.  
submitted to the EMBL Data Library, August 1997

A:Reference number: Z21825

A:Accession: T39072

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1687 <MUR>

A:Cross-references: UNIPROT:Q92355; EMBL:Z81317; PDB:QAB03612.1; GSPDB:GN00066; SPDB:SE

C:Experimental source: strain 972h; cosmid c6G9

C:Genetics:  
A:Gene: SPDB:SPAC6G9.010c

A:Map position: 1

Query Match 81.5%; Score 44; DB 2; Length 1687;  
Best Local Similarity 64.3%; Pred. No. 1.3;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGVITPYXXQV 14  
|||  
Db 1533 IGVITPYRQGV 1546

## RESULT 8

D72258  
helicase-related protein - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C:Accession: D72258

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.V.; Haft, D.H.; Hickey

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: D72258

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-245 <ARN>

A:Cross-references: UNIPROT:Q9X1D1; GB:AE001793; GB:AE000512; NID:94981963; PDB:ADJ3648

A:Experimental source: strain MSB8

C:Genetics:  
A:Gene: TM1411

Query Match 79.6%; Score 43; DB 2; Length 245;  
Best Local Similarity 72.7%; Pred. No. 0.26;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGVITPYXXQV 11  
|||  
Db 144 IGVITPYDDQV 154

RESULT 9  
D69085

transcription control factor enhancer-binding protein - Methanobacterium thermoautotroph  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C:Accession: D69085

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

Oliver, D.; Spadator, R.; Viscare, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivanti, N.

Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct

A:Reference number: A69000; MUID:98037514; PMID:93711463

A:Accession: D69085

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-642 <MTH>

A:Cross-references: UNIPROT:Q27671; GB:AE000922; GB:AE000666; NID:92622754; PDB:AAB8610

A:Experimental source: strain Delta H

C:Genetics:  
A:Gene: MTH1634

C:Superfamily: probable DNA helicase MJ0104

Query Match 79.6%; Score 43; DB 2; Length 642;  
Best Local Similarity 72.7%; Pred. No. 0.73;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGVITPYXXQV 11  
|||  
Db 532 IGVITPYDDQV 542

## RESULT 10

G72429  
hypothetical protein TM0005 - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C:Accession: G72429

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.V.; Haft, D.H.; Hickey

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: G72429

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-650 <ARN>

A:Cross-references: UNIPROT:Q9X1D1; GB:AE001689; GB:AE000512; NID:94980483; PDB:ADJ3509

A:Experimental source: strain MSB8

C:Genetics:  
A:Gene: TM0005

C:Superfamily: probable DNA helicase MJ0104

Query Match 79.6%; Score 43; DB 2; Length 650;  
Best Local Similarity 72.7%; Pred. No. 0.74;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGVITPYXXQV 11  
|||  
Db 549 IGVITPYDDQV 559

## RESULT 11

T52521  
related to SEM1 protein [imported] - Neurospora crassa

N:Alternate names: protein B2023.170

C:Species: Neurospora crassa

C:Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004

C:Accession: T52521

R:Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, September 2000

A:Reference number: Z26053

A:Accession: T52521

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1825 <SCH>  
 A:Cross-references: UNIPROT:Q9HRT5; EMBL:AL442164; GSPDB:GN00116; NCSP:B2J23.170  
 A:Experimental source: BAC clone B2J23; strain OR74A  
 C:Genetics:  
 A:Gene: NCSP:B2J23.170  
 A:Map position: 6

Query Match 79.6%; Score 43; DB 2; Length 1825;  
 Best Local Similarity 57.1%; Pred. No. 2.2;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IGVITPYXXQVXXL 14  
 ||:|||||:  
 Db 1558 IGVITPYKQQLAVL 1571

# RESULT 12

D86303  
 F17P16.1 protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: D86303  
 R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Hultzar, L.

Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzilli,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; PMID:11130712

A:Accession: D86303  
 A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-2142 <STO>

A:Cross-references: UNIPROT:Q9FWR3; GB:AE005172; NID:g9954728; PIDN:AA09081.1; GSPDB:GN  
 C:Genetics:  
 A:Map position: 1

Query Match 79.6%; Score 43; DB 2; Length 2142;  
 Best Local Similarity 57.1%; Pred. No. 2.7;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IGVITPYXXQVXXL 14  
 ||:|||||:  
 Db 1687 IGVITPYKQQLAVL 1700

# RESULT 13

D90303  
 hypothetical protein SSO1456 [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004

C:Accession: D90303

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyer, M.J.; Chan-  
 Jongs, I.; Jeffries, A.C.; Kozera, C.U.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.  
 arett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: D90303

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-139 <KDR>

A:Cross-references: UNIPROT:Q9YX82; GB:AE006641; NID:g13814681; PIDN:AAK41683.1; GSPDB:G  
 C:Genetics:  
 A:Gene: SSO1456

Query Match 77.8%; Score 42; DB 2; Length 139;  
 Best Local Similarity 57.1%; Pred. No. 0.23;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IGVITPYXXQVXXL 14  
 ||:|||||:  
 Db 52 VGVISPYRQVRL 65

# RESULT 14

D70476  
 DNA helicase - Aquifex aeolicus

C:Species: Aquifex aeolicus

C>Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004

C:Accession: D70476

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove  
 v.

Nature 392, 353-358, 1998  
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; PMID:9819666; PMID:9557320

A:Accession: D70476

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-530 <AGF>

A:Cross-references: UNIPROT:Q67840; GB:AE000770; NID:g2984274; PIDN:AA07803.1; PID:g2384  
 A:Experimental source: strain VFS

C:Genetics:  
 A:Gene: helX

Query Match 77.8%; Score 42; DB 2; Length 530;  
 Best Local Similarity 64.3%; Pred. No. 0.97;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IGVITPYXXQVXXL 14  
 ||:|||||:  
 Db 428 IGVISPYEDQVRL 441

# RESULT 15

F85069  
 hypothetical protein AT4G05540 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004

C:Accession: F85069

R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
 Nature 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; PMID:10617198

A:Accession: F85069

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-660 <STO>

A:Cross-references: UNIPROT:Q9S9W3; GB:NC\_001268; NID:g7267314; PIDN:CAB81096.1; GSPDB:GN  
 C:Genetics:  
 A:Gene: AT4G05540  
 A:Map position: 4

Query Match 77.8%; Score 42; DB 2; Length 660;  
 Best Local Similarity 50.0%; Pred. No. 1.2;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IGVITPYXXQVXXL 14  
 ||:|||||:  
 Db 565 VGVISPYKQVSAI 578

Search completed: April 18, 2005, 08:06:04  
 Job time : 13.4891 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: April 18, 2005, 07:33:55 ; Search time 58.4526 Seconds  
(without alignments)

122.648 Million cell updates/sec

Title: US-10-652-334-7  
Perfect score: 54  
Sequence: 1 IGVITPYXXQVXXL 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	87.0	635	Q9FNX9	Q9FNX9 arabidopsis
2	47	87.0	639	Q81047	Q81047 arabidopsis
3	47	87.0	651	Q6F2U8	Q6F2U8 oryza sativ
4	46	85.2	648	Q28883	Q28883 archaeoglob
5	46	85.2	1024	Q75U80	Q75U80 dictyostell
6	45	85.2	1944	Q94387	Q94387 schizosacch
7	45	83.3	1274	ARMI DROME	Q615K9 drosophila
8	44	81.5	536	Q9F6V0	Q9F6V0 arabidopsis
9	44	81.5	657	Q75CL7	Q75CL7 aebhya goss
10	44	81.5	681	Q6CMK3	Q6CMK3 kluyveromyc
11	44	81.5	683	YKB7_YEAST	P34243 saccharomyc
12	44	81.5	1002	SDB3_ARATH	Q85YD9 arabidopsis
13	44	81.5	1040	Q7PTC1	Q7PTC1 anopheles g
14	44	81.5	1187	M10L_MOUSE	Q99MVS mus musculu
15	44	81.5	1211	M10L_HUMAN	Q9BXT6 homo sapien
16	44	81.5	1215	Q8BZ26	Q8BZ26 mus musculu
17	44	81.5	1617	Q8CH13	Q8CH13 mus musculu
18	44	81.5	1687	SN1_SCHPO	Q92355 schizosacch
19	44	81.5	1942	HELZ_HUMAN	P42694 homo sapien
20	44	81.5	1965	Q6DFV5	Q6DFV5 mus musculu
21	43	79.6	245	Q9X1D1	Q9X1D1 thermotoga
22	43	79.6	383	Q7PVA7	Q7PVA7 anopheles g
23	43	79.6	642	Q27671	Q27671 methanobact
24	43	79.6	649	Q6CFH6	Q6CFH6 yarrowia li
25	43	79.6	650	Q9WXM0	Q9WXM0 thermotoga
26	43	79.6	695	Q6FKP9	Q6FKP9 candida gla
27	43	79.6	1361	Q8LNM4	Q8LNM4 oryza sativ
28	43	79.6	2126	Q9HPT5	Q9HPT5 neurospora
29	43	79.6	2142	Q9FMR3	Q9FMR3 arabidopsis
30	43	79.6	2646	Q7XPT5	Q7XPT5 oryza sativ
31	42	77.8	139	Q97Y82	Q97Y82 sulfolobus

32	42	77.8	530	2	Q67840	Q67840 aquifex ao
33	42	77.8	660	2	Q9S9W3	Q9S9W3 arabidopsis
34	42	77.8	849	2	Q8RDR3	Q8RDR3 fusobacteri
35	42	77.8	967	2	Q86AS0	Q86AS0 dictyostell
36	42	77.8	1121	1	Y806_YEAST	Y806_YEAST saccharomyc
37	42	77.8	1124	1	Q8A0E7	Q8A0E7 bacteroides
38	42	77.8	1153	2	Q64XY7	Q64XY7 bacteroides
39	41	75.9	338	1	Y0HR_ECOLI	Y0HR_ECOLI P33369 escherichia
40	41	75.9	338	2	Q9AL64	Q9AL64 shigella fl
41	41	75.9	563	2	Q8SVT3	Q8SVT3 encyphalit
42	41	75.9	653	2	Q74ZU0	Q74ZU0 aebhya goss
43	41	75.9	1075	2	Q9SHX6	Q9SHX6 arabidopsis
44	41	75.9	1076	2	Q9SHX8	Q9SHX8 arabidopsis
45	41	75.9	1090	2	Q64476	Q64476 arabidopsis

## ALIGNMENTS

RESULT 1						
ID	Q9FNX9	PRELIMINARY;	PRT;	635 AA.		
AC	Q9FNX9	01-MAR-2001 (TREMBlrel. 16, Created)				
DT	01-MAR-2001 (TREMBlrel. 16, Last sequence update)					
DT	01-MAR-2004 (TREMBlrel. 26, Last annotation update)					
DE	Purative helicase.					
CN	Name=atpc-2 gene;					
OS	Arabidopsis thaliana (Mouse-ear cress).					
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;					
OC	Eurosida II; Brassicales; Brassicaceae; Arabidopsids.					
OX	NCBI_TaxID=3702;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Whole plant;					
RA	Bezhani S., Sheremeti I., Pfannschmidt T., Oelmuejler R.;					
RT	"A repressor with similarities to pro- and eukaryotic DNA helicases					
RT	controls the assembly of the CAAT-box binding complex at a					
RT	photosynthesis gene promoter.";					
RL	J. Biol. Chem. 10:1074-1074(2001).					
DR	EMBL; AJ300306; CAC16347.1; -.					
DR	GO; GO:0005524; F:ATP binding; IEA.					
DR	GO; GO:0003677; F:DNA binding; IEA.					
DR	GO; GO:0004386; F:helicase activity; IEA.					
DR	GO; GO:0000166; F:nucleotide binding; IEA.					
DR	InterPro; IPR003593; AAA_Atpase.					
DR	InterPro; IPR011545; DEAD/DEAF N.					
DR	InterPro; IPR004483; put_DNA_helic.					
DR	SMART; SM00382; AAA; 1.					
DR	SMART; SM00487; DEXDC; 1.					
DR	TIGRPFAMs; TIGR00376; put_DNA_helic; 1.					
KW	ATP-binding; Helicase.					
SQ	SEQUENCE 635 AA; 70335 MW; 6149646686631F29 CRC64;					
Query Match 87.0%; Score 47; DB 2; Length 635;						
Best Local Similarity 64.3%; Pred. No. 0.61;						
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;						
Qy	1 IGVITPYXXQVXXL 14					
Db	524 IGVITPYXXQVXXL 537					
RESULT 2						
AC	Q81047	PRELIMINARY;	PRT;	639 AA.		
ID	Q81047	094AT0;				
DT	01-NOV-1998 (TREMBlrel. 08, Created)					
DT	01-JUN-2002 (TREMBlrel. 21, Last sequence update)					
DT	05-JUN-2004 (TREMBlrel. 27, Last annotation update)					
DE	Purative helicase.					
GN	Name=At2g03270;					

OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eustoids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., Vanaken S.E.,  
 RA Bannstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Bantico M.-I.,  
 RA Carreira A.J., Creasy T.H., Buell C.R., Town C.D., Nieman W.C.,  
 RA Fraser C.M., Venter J.C.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Town C.D., Kaul S.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,  
 RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
 RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi Y., Yu G., Bowser L.,  
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ichida J., Jones T.,  
 RA Kamita A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,  
 RA Seki M., Shim P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
 RA Theologis A.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,  
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,  
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,  
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ichida J., Jones T.,  
 RA Kamita A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,  
 RA Sakurai T., Satou M., Seki M., Shim P., Southwick A., Shinozaki K.,  
 RA Davis R.W., Ecker J.R., Theologis A.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC006284; AADI7447.2; -  
 DR EMBL: AC005313; AAM15033.1; -  
 DR EMBL: AY045820; AAT6484.1; -  
 DR EMBL: AY091361; AAM14300.1; -  
 DR PIR: T02699; T02699.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR GO: GO:0004386; F:helicase activity; IEA.  
 DR GO: GO:0000166; F:nucleotide binding; IEA.  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR011545; DEAD/DEAH\_N.  
 DR InterPro: IPR004483; put\_DNA\_helic.  
 DR SMART: SM00382; AAA; 1.  
 DR SMART: SM00487; DEXDC; 1.  
 DR TIGRFAMs: TIGR00376; put\_DNA\_helic; 1.  
 DR ATP-binding; Helicase.  
 KW K  
 SQ SEQUENCE 639 AA; 70850 MW; C91340DAE4C978C CRC64;  
 Query Match 87.0%; Score 47; DB 2; Length 639;  
 Best Local Similarity 64.3%; Pred. No. 0.62;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

DT 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE Putative DNA helicase.  
 GN Name=OSUNBA0010D22.27;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriocarpaceae; Oryzae; Oryza.  
 NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,  
 RA Overton D.W., Tallon L.J., Kim M.M., Bera J.J., Jin S.S.,  
 RA Padrosh D.M., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,  
 RA Vanaken S.S., Riedmiller S.B., Uteback T.T., Feldblyum T.V.,  
 RA Yang Q.Q., Haas B.U., Suh B.B., Peterson J.J., Quackenbush J.,  
 RA White O., Salzberg S.L., Fraser C.M.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Buell R.;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC133003; AAT76339.1; -  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR GO: GO:0004386; F:helicase activity; IEA.  
 DR GO: GO:0000166; F:nucleotide binding; IEA.  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR011545; DEAD/DEAH\_N.  
 DR InterPro: IPR004483; put\_DNA\_helic.  
 DR SMART: SM00382; AAA; 1.  
 DR SMART: SM00487; DEXDC; 1.  
 DR TIGRFAMs: TIGR00376; put\_DNA\_helic; 1.  
 DR ATP-binding; Helicase.  
 KW K  
 SQ SEQUENCE 651 AA; 71304 MW; EC701B00C458ED9 CRC64;  
 Query Match 87.0%; Score 47; DB 2; Length 651;  
 Best Local Similarity 64.3%; Pred. No. 0.63;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

DT 25-OCT-2004 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE DNA helicase, putative.  
 GN OrderedLocusNames=AF1388;  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
 OC Archaeoglobaceae; Archaeoglobus.  
 NCBI\_TaxID=2234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STEAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RC MEDLINE=98049343; PubMed=9389475; DOI=10.1038/37052;  
 RA Klein H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 RA Ketchum K.A., Dodson R.J., Gwinn M.L., Hickey E.R., Peterson J.D.,  
 RA Richardson D.L., Kerlavage A.R., Graham D.B., Kyriakides N.C.,  
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.R.,  
 RA Kirches E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.J.,  
 RA Peterson S.N., Reich C.I., McNeill L.K., Badger J.H., Glodek A.,  
 RA Zhou U., Overbeek R., Goehne J.D., Weidman J.F., McDonald L.A.,  
 RA Uteback T.R., Cotton M.D., Spitzgr T., Artlich P., Kaine B.P.,  
 RA Sykes S.M., Sadow P.W., D'Andrea K.P., Bowman C., Fujii C.,  
 RA Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O.,

RA Woese C.R., Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 reducing archaeon *Archaeoglobus fulgidus*.";  
 RL Nature 390:364-370(1997).  
 DR EMBL; AE001009; AAB89860.1; -.  
 DR PIR; C69423; C69423.  
 DR TIGR; AF1388; -.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0004386; F:helicase activity; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR InterPro; IPR003593; AAA\_Arpase.  
 DR InterPro; IPR004483; put\_DNA\_helic.  
 DR SMART; SM00382; AAA; 1.  
 DR TIGRPMs; TIGR00376; put\_DNA\_helic; 1.  
 KM ATP-binding; Complete proteome; Helicase.  
 SQ SEQUENCE 648 AA; 73775 MW; B5074507D856A7D5 CRC64;

Query Match 85.2%; Score 46; DB 2; Length 648;  
 Best Local Similarity 71.4%; Pred. No. 1;  
 Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGVTPYXXQVXXL 14  
 DB 537 IGVTPYDDQVDDL 550

RESULT 5  
 ID Q75JAO PRELIMINARY; PRT; 1024 AA.  
 AC Q75JAO;  
 DT 05-JUN-2004 (TREMblrel. 27, Created)  
 DT 05-JUN-2004 (TREMblrel. 27, Last annotation update)  
 DT 05-JUN-2004 (TREMblrel. 27, Last annotation update)  
 DE Similar to Rattus norvegicus (Rat). Antifreeze-enhancer binding  
 DE protein AEP.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.  
 ON NCBI\_TaxId=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AA4;  
 RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;  
 RA Gloeckner G., Richter L., Szafranski K., Pachebat J., Dear P.,  
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,  
 RA Tunggal B., Cox B., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;  
 RT "Sequence and analysis of chromosome 2 of *Dictyostelium discoideum*.";  
 RL Nature 418:79-85(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AA4;  
 RA Baumgart C.;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC117072; AAS38931.1; -.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR InterPro; IPR003593; AAA\_Arpase.  
 DR InterPro; IPR011545; DEAD/DEAF N.  
 DR InterPro; IPR004483; put\_DNA\_helic.  
 DR InterPro; IPR001374; RSH\_  
 DR InterPro; IPR000058; Znf\_AN1.  
 DR Pfam; PF01424; RSH; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00393; RSH; 1.  
 DR SMART; SM00154; Znf\_AN1; 1.  
 DR TIGRPMs; TIGR00376; put\_DNA\_helic; 1.  
 KM ATP-binding.  
 SQ SEQUENCE 1024 AA; 115697 MW; 92352F4C44885C5A CRC64;

Query Match 85.2%; Score 46; DB 2; Length 1024;  
 Best Local Similarity 71.4%; Pred. No. 1.5;

Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGVTPYXXQVXXL 14  
 DB 586 IGVTPYNGQVXL 599

RESULT 6  
 ID Q94387 PRELIMINARY; PRT; 1944 AA.  
 AC Q94387;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)  
 DE SPBC29A10.10c protein.  
 GN Name=SPBC29A10.10c;  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 ON NCBI\_TaxId=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972h-;  
 RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Mablett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovich E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S., Warren T., Whitehead S.,  
 RA Woodward J., Voiclaert G., Aert R., Robben J., Grymoprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaffer M., Muller-Auer S.,  
 RA Gabel C., Fuchs M., Duetschke A., Fritz C., Hoier E., Moestl D.,  
 RA Hilbert H., Borzym K., Langer I., Beck A., Leinach H., Reinhardt R.,  
 RA Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaue V., Motier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Roche M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of *Schizosaccharomyces pombe*.";  
 RL Nature 415:871-880(2002).  
 DR EMBL; AL034463; CAA22438.1; -.  
 DR PIR; T40065; T40065.  
 DR GenDB SPombe; SPBC29A10.10c; -.  
 DR InterPro; IPR000873; AMP-Bind.  
 DR PROSITE; PS00455; AMP BINDING; UNKNOWN 1..  
 SQ SEQUENCE 1944 AA; 222209 MW; 12B00593ABEB11C CRC64;

Query Match 85.2%; Score 46; DB 2; Length 1944;  
 Best Local Similarity 64.3%; Pred. No. 2.8;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGVTPYXXQVXXL 14  
 DB 1671 IGVTPYRSQVQOL 1684

RESULT 7  
 ID ARMT\_DROME STANDARD; PRT; 1274 AA.  
 AC Q6J5K9; Q9VZP4;  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)



DT 25-JAN-2005 (Rel. 46, Last annotation update)  
 DE Probable RNA helicase arm1 (EC 3.6.1.-) (Armitage protein).  
 GN Name=arm1; ORFNames=CG11513.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI Taxid=7227;  
 [1]  
 RP SEQUENCE FROM N.A. (ISOFORM B), FUNCTION, SUBCELLULAR LOCATION, AND  
 RP TISSUE SPECIFICITY.  
 RC TISSUE=Ovary.  
 RX PubMed=15035984; DOI=10.1016/S0092-8674(04)00250-8;  
 RA Cook H.A., Koppetsch B.S., Wu J., Theurkauf W.E.;  
 RT "The Drosophila SDE3 homolog arm1age is required for oskar mRNA  
 RT silencing and embryonic axis specification.";  
 RL Cell 116:817-829(2004).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.P.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blake J.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abil U.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.W., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolashkov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brockett P.,  
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,  
 RA Folsler G., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iodgwan C.,  
 RA Kamil M.B., Kaulash F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Palazotto K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Skupski M.P., Smith T.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zheng H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 [3]  
 RP GENOME REANNOTATION.  
 RP MEDLINE=22426069; PubMed=12537572;  
 RA Mera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.G., Bayraktaroglu L., Berman B.P.,  
 RA Betencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 [4]

RP FUNCTION.  
 RX PubMed=15035984; DOI=10.1016/S0092-8674(04)00218-1;  
 RA Tomari Y., Du T., Haley B., Schwarz D.S., Bennett R., Cook H.A.,  
 RA Koppetsch B.S., Theurkauf W.E., Zamore P.D.;  
 RT "RISC assembly defects in the Drosophila RNAi mutant arm1age.";  
 RL Cell 116:831-841(2004).  
 CC -1- FUNCTION: Probable RNA helicase required for axial polarization of  
 CC the oocyte during early and mid oogenesis. Plays a central role in  
 CC RNA interference (RNAi) process, a process that mediates mRNA  
 CC destruction of translational repression. Required for the assembly  
 CC of the RISC complex, a complex required for target RNA destruction  
 CC or repression. May be required in the RISC assembly to unwind  
 CC mRNAs, in the production of single-stranded mRNA from the  
 CC double-stranded miRNA, a key step in RISC formation. Required both  
 CC for the translational control of oskar (osk) mRNA and cytoskeletal  
 CC polarization in the oocyte.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named Isoforms=2;  
 CC Name=A;  
 CC IsoId=O6J5K9-1; Sequence=Displayed;  
 CC Note=No experimental confirmation available;  
 CC Name=B;  
 CC IsoId=O6J5K9-2; Sequence=VSP\_011795, VSP\_011796;  
 CC -1- TISSUE SPECIFICITY: Abundant in oocytes and syncytial blastoderm.  
 CC Expressed at low level throughout development, including somatic  
 CC tissues. First apparent early in oogenesis, in the cytoplasm of  
 CC stem cells and mitotically dividing cystoblasts. In regions 2a and  
 CC 2b of the germarium, it is most concentrated in the center of the  
 CC germline cysts, where the pro-oocyte is located. In stage 1 and  
 CC oocyte, near the ring canals, it accumulates at the anterior of the  
 CC oocyte, near the ring canals. It also extends through the ring  
 CC canals forming a branched structure that links the early oocyte  
 CC with adjacent nurse cells. In stage 3 cysts, it accumulates at the  
 CC posterior cortex and localizes to extensions that pass through the  
 CC oocyte into the nurse cells. Through stages 4 to 7, it continues  
 CC to be somewhat enriched at the posterior cortex of the oocyte, but  
 CC at significantly lower level. In stage 9 to 10 egg chambers, it is  
 CC found throughout the cytoplasm of the oocyte and nurse cells, with  
 CC slight enrichment at the oocyte cortex.  
 CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family. SDE3  
 CC subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AY598469; AAT12000.1; -;  
 CC EMBL: AE003478; AAF47775.1; -;  
 CC DR FlyBase; FBgn0041164; arm1.  
 CC DR FlyBase; FBgn0035448; CG11513.  
 CC KW Alternative splicing; ATP-binding; Developmental protein; Helicase;  
 CC Hydrolyase.  
 CC NP BIND 809 816 ATP (potential).  
 CC SITE 948 951 DENG box (by similarity).  
 CC VARSPLIC 1 86 Missing (in isoform B).  
 CC VARSPLIC 87 92 /FTRD=VSP\_011795.  
 CC VARSPLIC 87 92 /FTRD=VSP\_011796.  
 CC FT CONFLICT 608 608 T -> K (in Ref. 1).  
 CC FT CONFLICT 613 613 I -> V (in Ref. 1).  
 CC FT CONFLICT 991 991 S -> L (in Ref. 1).  
 CC SQ SEQUENCE 1274 AA; 144887 MW; C340E2B584F75653 CRC64;  
 CC Query Match 83.3%; Score 45; DB 1; Length 1274;  
 CC Best Local Similarity 64.3%; Pred. No. 3;  
 CC Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 CC 1 IGVITPXXOVXXL 14



Db 1131 IGVITPYXQVXXL 1144

## RESULT 8

Q9FGV0 PRELIMINARY; PRT; 536 AA.

AC 09FGV0; 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE Gb|AAD48967.1.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 CC eurosid II; Brassicales; Brassicaceae; Arabidopsia.  
 CC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20181125; PubMed=10718197;  
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asanizu E., Kotani H.,  
 RA Tabata S.,  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC  
 RT clones."  
 RL DNA Ref. 7:31-63(2000).  
 DR EMBL: AB024026; BAB08973.1;  
 SQ SEQUENCE 536 AA; 60456 MW; 007355F5E527D89C CRC64;

Query Match 81.5%; Score 44; DB 2; Length 536;  
 Best Local Similarity 57.1%; Pred. No. 2.2;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IGVITPYXQVXXL 14  
 Db 409 VGVITPYXQVRAI 422

## RESULT 9

Q75CL7 PRELIMINARY; PRT; 657 AA.

AC 075CL7; 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE ACU098CP.  
 GN ORFNames=ACU098C;  
 OS Asbhyia gossypii (Yeast) (Eremothecium gossypii).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Eremothecium.  
 CC NCBI\_TaxID=33169;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 10895;  
 RA Brachet S., Voegelé S.E., Dietrich F.S., Lerch A., Gaffney T.,  
 RA Phillipsen P.;  
 RL Submitted (SEP-2004) to the EMBL/Genbank/DBJ databases.

DR EMBL: AE016816; AAS51130.1; -.  
 DR AGD; ACU098C; -.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR GO: GO:0000166; F:nucleotide binding; IEA.  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR01545; DEAD/DEAH\_N.  
 DR InterPro: IPR004483; put DNA helic.  
 DR Pfam: PF00580; UvrD-helicase; 1.  
 DR SMART: SM00383; AAA; 1.  
 DR SMART: SM00487; DEXDC; 1.  
 DR TIGRFAMs: TIGR00376; put\_DNA\_helic; 1.  
 KW ATP-binding.

SEQUENCE 657 AA; 72946 MW; 28C1905117D6DC2A CRC64;

Query Match 81.5%; Score 44; DB 2; Length 657;  
 Best Local Similarity 64.3%; Pred. No. 2.6;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IGVITPYXQVXXL 14  
 Db 545 IGVISPYNAQVALL 558

## RESULT 10

Q6CMS3 PRELIMINARY; PRT; 681 AA.

AC Q6CMS3; 25-OCT-2004 (TREMBLrel. 28, Created)  
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)  
 DE Similar to sp|P34243 Saccharomyces cerevisiae YKL017c D1P1 DNA  
 DE helicase A.  
 GN ORFNames=K1LA0E181509;  
 OS Kluyveromyces lactis NRRL Y-1140.  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.  
 CC NCBI\_TaxID=284390;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL Y-1140;  
 RG Genolevures;  
 RA Dujon B., Sherman D., Fischer G., Durans P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,  
 RA Goffard N., Frangeul L., Migle M., Anthouard V., Babour A., Barbe V.,  
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Bolsenne A., Boyer J., Cattolico L., Confalonieri F., de Darvar A.,  
 RA Despons L., Fabre B., Fairhead C., Ferry-Dumazet H., Groppi A.,  
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicoud J.M., Nikolski M., Oztas S., Ozler-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
 RA Swenne D., Tekela F., Wesolowski-Louvel M., Westhof B., Wirth B.,  
 RA Zentou-Meyer M., Zivanovic I., Bolotin-Pukhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,  
 RA Wincker P., Souciet J.L.;  
 RT "Genome evolution in yeasts."  
 RL Nature 430:35-44(2004).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL Y-1140;  
 RG Genoscope;  
 RA Submitted (JUL-2004) to the EMBL/Genbank/DBJ databases.  
 RL EMBL: CR382125; CAG99853.1; -.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR GO: GO:0004386; F:helicase activity; IEA.  
 DR GO: GO:0000166; F:nucleotide binding; IEA.  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR01545; DEAD/DEAH\_N.  
 DR InterPro: IPR004483; put DNA helic.  
 DR InterPro: IPR000212; UvrD-helicase.  
 DR Pfam: PF00580; UvrD-helicase; 1.  
 DR SMART: SM00382; AAA; 1.  
 DR SMART: SM00487; DEXDC; 1.  
 DR TIGRFAMs: TIGR00376; put\_DNA\_helic; 1.  
 KW ATP-binding; Helicase.

SEQUENCE 681 AA; 77155 MW; D3878209DCE6377 CRC64;

Query Match 81.5%; Score 44; DB 2; Length 681;  
 Best Local Similarity 64.3%; Pred. No. 2.7;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IGVITPYXQVXXL 14  
 Db 565 IGVISPYNSQVSL 578

```

RESULT 11
ID YK87 YEAST STANDARD; PRT; 683 AA.
AC P34243
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Hypothetical 78.3 kDa protein in RAM2-ATP7 intergenic region.
GN OrderedlocusNames=YKL017C;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94205264; PubMed=8154185;
RA Wiemann S., Voss H., Schwager C., Rupp T., Stegemann J.,
RA Zimmermann J., Grothues D., Sensen C., Erle H., Hewlett N.,
RA Banerji A., Ansorge W.;
RT "Sequencing and analysis of 51.6 kilobases on the left arm of
RT chromosome XI from Saccharomyces cerevisiae reveals 23 open reading
RT frames including the PMS1 gene."
RL Yeast 9:1343-1348(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Rieger M.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
DR EMBL; X74152; CA52266.1; -
DR EMBL; Z28017; CA81852.1; -
DR PIR; S34700; S34700.
DR Geronlone; 139773; -.
DR SGD; S000001500; HCS1.
DR GO; GO:0005658; C:alpha DNA polymerase:primase complex; IPI.
DR GO; GO:0003678; F:DNA helicase activity; IDA.
DR GO; GO:0006273; P:laging strand elongation; IPI.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR004483; put_DNA_helic.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00487; DEXDC; 1.
DR TIGRfams; TIGR00376; put DNA helic; 1.
KW ATP-binding; Helicase; Hypothetical protein.
FT NP BIND 229 236 ATP (Potential).
SQ SEQUENCE 683 AA; 78272 MW; FCD026807573778 CRC64;
OY 1 IGVITPYXXQVXXL 14
Db 569 IGVISPYNAQVSHL 582
Query Match 81.5%; Score 44; DB 1; Length 683;
Best Local Similarity 64.3%; Pred. No. 2.7;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

RESULT 12
SDE3_ARATH STANDARD; PRT; 1002 AA.
ID SDE3_ARATH
AC O8GYD9; O94K51; O9LR48;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Probable RNA helicase SDE3 (EC 3.6.1.-) (Silencing defective protein
DE 3).
GN Name=SDE3; OrderedlocusNames=At1g05460; ORFNames=T25N20.11, T25N20.26;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=21192042; PubMed=11296239; DOI=10.1093/emboj/20.8.2069;
RA Dalmay T., Horsfield R., Braunstein T.H., Bailcombe D.C.;
RT "SDE3 encodes an RNA helicase required for post-transcriptional gene
RT silencing in Arabidopsis."
RL EMBO J. 20:2069-2078(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RP MEDLINE=21016719; PubMed=11130712; DOI=10.1038/35048500;
RA Theologis A., Ecker J.R., Palm C.J., Federici N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.U., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.-J., Koo H.L., Kremetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Malti R., Matzli A.,
RA Millicher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tumbunga G., Toriumi M.J., Town C.D.,
RA Unterhark T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RX PubMed=11910074; DOI=10.1126/science.1071006;
RA Seki M., Narusaka M., Kamiya A., Ishida J., Satou M., Sakurai T.,
RA Nakajima M., Enji A., Akiyama K., Oono Y., Muramatsu M.,
RA Hayashizaki Y., Kawai Y., Carninci P., Itoh M., Ishii Y., Arakawa T.,
RA Shibata K., Shingawa A., Shinozaki K.;
RT "Functional annotation of a full-length Arabidopsis cDNA collection."
RL Science 296:141-145(2002).
RN [4]
RP FUNCTION.
RX PubMed=12941703; DOI=10.1093/emboj/cdg431;
RA Humber C., Dunoyer P., Moissard G., Ritzenthaler C., Voimnet O.;
RT "Transitivity-dependent and -independent cell-to-cell movement of RNA
RT silencing."
RL EMBO J. 22:4523-4533(2003).
CC -1- FUNCTION: Probable RNA helicase required for post-transcriptional
CC gene silencing (PTS), a process that provides protection in
CC plants against virus infection and can suppress expression of
CC transgenes. Plays a central role in RNA interference (RNAi)
CC process, a process that mediates mRNA destruction of translational
CC repression. Required for the assembly of the RISC complex, a
CC complex required for target RNA destruction or repression. May be
CC required in the RISC assembly to unwind mRNAs, in the production
CC of single-stranded mRNA from the double-stranded mRNA, a key
CC step in RISC formation.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family. SDE3
CC subfamily.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; AF39908; AAK40099.1; -.  
DR EMBL; AC005106; AAF79736.1; ALT\_SEQ.  
DR EMBL; AK117698; BAC42350.1; -.  
DR PIR; C6189; C6189.  
KM ATP-binding; Helicase; Hydrolase.  
FT NP\_BIND 421 428 ATP (potential).  
FT SITE 545 548 DEAG box.  
FT CONFLICT 329 329 M -> L (in Ref. 1).  
FT CONFLICT 535 535 N -> K (in Ref. 1).  
FT CONFLICT 646 646 D -> E (in Ref. 1).  
FT CONFLICT 991 991 A -> G (in Ref. 1).  
SQ SEQUENCE 1002 AA; 113362 MW; 32D906F7FA95C0A5 CRC64;  
Query Match 81.5%; Score 44; DB 1; Length 1002;  
Best Local Similarity 81.8%; Pred. No. 3.9;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGVITPYXXQV 11  
Db 709 IGVITPYRQGV 719

RESULT 13  
Q7PTC1 PRELIMINARY; PRT; 1040 AA.  
ID 07PTC1  
AC 07PTC1; 26. Created  
DT 01-MAR-2004 (TREMBL) 26. Last sequence update  
DT 01-MAR-2004 (TREMBL) 26. Last annotation update  
DE ENSANGP0000021787 (fragment).  
GN Name=ENSANGP0000021787 (fragment).  
OS Anopheles gambiae str. PEST.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
OX NCBI\_TaxID=16045;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PEST;  
RA Anopheles Genome Sequencing Consortium;  
RL Submitted (APR-2003) to the EMBL/Genbank/DBJ databases.  
CC -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AAB0100807; EAA04125.2; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 1040 AA; 116035 MW; 53478A659989E96 CRC64;  
Query Match 81.5%; Score 44; DB 2; Length 1040;  
Best Local Similarity 57.1%; Pred. No. 4.1;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGVITPYXXQV 14  
Db 931 IGVITPYMQVRSI 944

RESULT 14  
M10L\_MOUSE STANDARD; PRT; 1187 AA.  
ID M10L\_MOUSE  
AC Q93MV5; Q7TPA9; Q8C3W0; Q924C2;  
DT 28-FEB-2003 (Ref. 41, Created)  
DT 28-FEB-2003 (Ref. 41, Last sequence update)  
DT 25-OCT-2004 (Ref. 45, Last annotation update)  
DE Potential helicase Mov10l1 (EC 3.6.1.-) (Mooney leukemia virus 10-  
DE like protein 1) (MOV10-like 1) (Cardiac helicase activated by MEF2  
DE protein) (Cardiac-specific RNA helicase).  
GN Name=Mov10l1; Synonyms=Champ;  
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Testis;  
RX MEDLINE=21175748; PubMed=11279525; DOI=10.1038/86927;  
RA Wang P.J., McCarrey J.R., Yang F., Page D.C.;  
RT "An abundance of X-linked genes expressed in spermatogonia.";  
RL Nat. Genet. 27:422-426(2001).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RC STRAIN=NIH Swiss; TISSUE=Heart;  
RX MEDLINE=12590829; PubMed=11397016; DOI=10.1006/dbio.2001.0277;  
RA Liu Z.-P., Nakagawa O., Nakagawa M., Yanagisawa H., Passler R.,  
RA Richardson J.A., Sivasubra D., Olson E.N.;  
RT "ChAMP, a novel cardiac-specific helicase regulated by MEF2C";  
RL Dev. Biol. 224:497-509(2001).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 3), INDUCTION, AND TISSUE SPECIFICITY.  
RC STRAIN=NIH Swiss;  
RX PubMed=12754203; DOI=10.1074/jbc.M300014200;  
RA Ueyama T., Kasahara H., Ishiwata T., Yamashita N., Izumo S.;  
RT "Cam, a cardiac-specific isoform of the RNA helicase Mov10l1, is  
RT regulated by Nkx2.5 in embryonic heart.";  
RL J. Biol. Chem. 278:28750-28757(2003).  
RN [4]  
RP SEQUENCE OF 769-1187 FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Heart;  
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nikiel D., Otsu N., Saito R., Suzuki H., Yamashita I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schombach C., Gojobori T.,  
RA Bairdrelli R., Hill D.P., Bult C., Hume D.A., Quakenbush J.,  
RA Schiraldi L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.N.,  
RA Blake J.A., Bradt D., Bruscia V., Chochia C., Corbani L.B., Cousins S.,  
RA Dalla E., Dregani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi S., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Guerinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numa K., Okido T., Pavan W.J., Petrea G., Pasole G.,  
RA Petrovsky N., Piliat R., Pontius J.U., Qi D., Ramchandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sadelain A., Schneider C., Semple C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wysshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [5]  
RP FUNCTION OF ISOFORM 2, AND MUTAGENESIS OF 888-ASP--GLN-892.  
RX PubMed=11854500; DOI=10.1073/pnas.261708699;  
RA Liu Z.-P., Olson E.N.;  
RT "Suppression of proliferation and cardiomyocyte hypertrophy by CHAMP,  
RT a cardiac-specific RNA helicase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:2043-2048(2002).  
CC -1- FUNCTION: Putative RNA helicase. Isoform 1 may play a role in male  
CC germ cell development and isoform 2 may act downstream of MEF2C  
CC during heart formation. Isoform 2 acts as a cardiac-specific  
CC suppressor of cardiomyocyte hypertrophy and cell cycle  
CC progression, suggesting that it may suppress these processes  
CC through the regulation of CDKN1A.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (isoform 2).



RA Martyn I.D., Maehreghi-Mohammadi M., Matthews L.H., Mccann O.T.,  
RA Mcclay J., McLaren S., McMurray A.V., Milne S.A., Mortimore B.J.C.T.,  
RA Odell C.N., Pavlitt R., Pearce A.A., Pearson D., Phillips B.J.C.T.,  
RA Phillips S.H., Plund R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,  
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,  
RA Soderlund C., Spurgeon L., Steward C.A., Sulston J.E., Swann R.M.,  
RA Vaudin M., Wall M., Wallie J.M., Whiteley M.N., Willey D.L.,  
RA Williams L., Williams S.A., Williamson H., Wilner T.B., Winding L.,  
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,  
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyma S.,  
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do H., Do T.,  
RA Dorman A., Pang F., Hu Y., Hua A., Kenton S., Lai H., Lao H.I.,  
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,  
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shauli S., Sloan D., Song L.,  
RA Wang Q., Wang Y., Wang Z., White J., Williamson D., Wu H., Yao Z.,  
RA Zhan M., Zhang G., Chisase S., Murray J., Miller N., Min P.,  
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,  
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawking J.,  
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,  
RA Schaefer P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,  
RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,  
RA Wilson R., Emanuel B.S., Shaikh T., Kuraishi H., Salta S.,  
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.B.,  
RA Edelmann L., Kim U.J., Shikuya H., Simon M.I., Dunanek J.P.,  
RA Peyrard M., Kedra D., Serousi E., Fransson I., Tapia I., Bruder C.E.,  
RA O'Brien K.P., Wilkison P., Bodenteich A., Hartman K., Hu X.,  
RA Khan A.S., Lane L., Tikhunov Y., Wright H.,  
RA "The DNA sequence of human chromosome 22." ;  
RA Nature 402:489-495 (1999).  
RN [4]  
RP SEQUENCE OF 534-1211 FROM N.A., AND VARIANTS LEU-454; ARG-820 AND  
RP GUD-1179.  
RC TISSUE=Testis;  
RA Bloecher H., Boecher W., Brandt P., Mewes H.-W., Gassenhuber J.,  
RA Wiemann S.;  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Putative RNA helicase. Isoform 1 may play a role in male  
CC germ cell development.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Comment=Additional isoforms seem to exist;  
CC Name=1;  
CC IsoId=Q9BXT6-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q9BXT6-2; Sequence=VSP\_003390, VSP\_003391;  
CC Name=3;  
CC IsoId=Q9BXT6-3; Sequence=VSP\_010945, VSP\_010946, VSP\_010947,  
CC VSP\_010948;  
CC Note=No experimental confirmation available;  
CC -1- TISSUE SPECIFICITY: Isoform 1 is specifically expressed in testis.  
CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family. SDB3  
CC subfamily.  
CC -1- CAUTION: Ref.3 sequence differs from that shown due to erroneous  
CC gene model prediction.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF285604; AAK31983.1; -;  
CC EMBL; AK000033; BAA30895.1; -;  
CC EMBL; AK090740; BAC03511.1; -;  
CC EMBL; AL034546; -; NOT ANNOTATED CDS.  
CC EMBL; AL022328; CAB63041.1; AUT\_SEQ.  
CC EMBL; AL133068; CAB61391.1; -;  
CC PIR; T42668; T42668.  
CC Genew; HGNC:7201; MOV10L1.  
CC MIM; 605794; -.

DR GO; GO:0005622; C:intracellular; IC.  
DR GO; GO:0005524; F:ATP binding; TAS.  
DR GO; GO:0004004; F:ATP-dependent RNA helicase activity; NAS.  
DR GO; GO:0000287; F:magnesium ion binding; TAS.  
DR GO; GO:0003723; F:RNA binding; TAS.  
DR GO; GO:0007281; F:germ-cell development; IEI.  
DR GO; GO:0007283; F:spermatogenesis; IEI.  
DR InterPro; IPR008994; Nucleic acid DB.  
KW Alternative splicing; ATP-binding; Developmental protein; Helicase;  
KW Hydrolyase; Polymorphism.  
FT NP\_BIND 770 777 ATP (Potential).  
FT SITE 886 889 DEAD box.  
FT VARSPLIC 1 873 Missing (in isoform 2).  
FT VARSPLIC 874 875 GV -> MF (in isoform 2).  
FT VARSPLIC 1 902 Missing (in isoform 3).  
FT VARSPLIC 910 964 Missing (in isoform 3).  
FT VARSPLIC 1074 1082 Missing (in isoform 3).  
FT VARSPLIC 1083 1211 Missing (in isoform 3).  
FT VARIANT 454 454 I -> L (in dbSNP:2272843).  
FT VARIANT 650 650 /Frid=VAR\_013694.  
FT VARIANT 820 820 /Frid=VAR\_013695.  
FT VARIANT 1179 1179 Q -> R (in dbSNP:2272837).  
FT VARIANT 1063 1063 /Frid=VAR\_013696.  
FT CONFLICT 1096 1096 A -> E (in dbSNP:2272843).  
FT SEQUENCE 1211 AA; 135292 MW; 410716BDC85F8B1 CRC64;  
Query Match 81.5%; Score 44; DB 1; Length 1211;  
Best Local Similarity 81.8%; Pred. No. 4.7;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 IGVTPYXXOV 11  
DB 1063 IGVTPYRQOV 1073

Search completed: April 18, 2005, 08:15:47  
Job time : 59.4526 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: April 18, 2005, 07:10:15 ; Search time 114.496 Seconds  
(without alignments)  
77.692 Million cell updates/sec

Title: US-10-652-334-8

Perfect score: 103  
Sequence: 1 VEVXTVDXFGQREKDXIIISCVR 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 sege, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: geneseqp1980s:.\*  
2: geneseqp1990s:.\*  
3: geneseqp2000s:.\*  
4: geneseqp2001s:.\*  
5: geneseqp2002s:.\*  
6: geneseqp2003as:.\*  
7: geneseqp2003bs:.\*  
8: geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	94.2	23	7	ABW01184 Saccharom
2	97	94.2	23	8	ADP44110 Yeast tra
3	97	94.2	24	3	AAV77811 Mottif VII
4	95	92.2	380	3	AAV77814 Yeast Upf
5	95	92.2	380	7	ABW01205 Saccharom
6	95	92.2	380	8	ADP44131 Yeast hel
7	95	92.2	925	8	ADN19434 Bacteri
8	95	92.2	925	3	AAV98057 Yeast Upf
9	95	92.2	971	6	ABR53412 Protein s
10	95	92.2	971	7	ADK64706 Diasease c
11	94	91.3	373	8	ADRI0398 Human pro
12	94	91.3	797	2	AAW36508 Human REN
13	94	91.3	935	7	ADJ69900 Human hea
14	94	91.3	992	8	ADS23757 Bacteri
15	94	91.3	1043	2	AAW36509 Murine RE
16	94	91.3	1140	2	AAV21377 Human HUP
17	89	86.4	1118	6	ABG73900 Human REN
18	86	83.5	414	3	AAV77816 Yeast Dip
19	86	83.5	414	7	ABW01204 Saccharom
20	86	83.5	414	8	ADP44130 Yeast hel
21	86	83.5	562	7	ADB63849 Human pro
22	86	83.5	677	6	ABR82240 Human nuc
23	86	83.5	683	8	ADS43834 Bacteri
24	86	83.5	926	8	ADQ66763 Novel hum
25	84	81.6	818	8	ADK68058 Female re

26	83	80.6	324	4	AAB94335	Aab94335 Human pro
27	83	80.6	828	4	AAW39674	Aam39674 Human pol
28	83	80.6	828	4	AAB92977	Aab92977 Human pro
29	83	80.6	828	6	ABR56327	AbR56327 NM 015046
30	83	80.6	828	8	ADP23318	Adp23318 PRO poly
31	83	80.6	829	4	AAW80279	Aam80279 Human pro
32	83	80.6	829	4	AAW41460	Aam41460 Human pro
33	83	80.6	966	4	AAW79295	Aam79295 Human pro
34	83	80.6	2677	6	ABR58288	AbR58288 BCT0156 P
35	82	79.6	1944	8	ADN19868	Adn19868 Bacteri
36	81	78.6	642	8	ADS43096	AdS43096 Bacteri
37	80	77.7	653	8	ADS43215	AdS43215 Bacteri
38	80	77.7	654	4	AAB96566	Aab96566 Pucative
39	79	76.7	648	8	ADS41904	AdS41904 Bacteri
40	79	76.7	655	4	AAB62031	Aab62031 Recombina
41	79	76.7	656	8	ADN18615	Adn18615 Bacteri
42	79	76.7	698	7	ADM25464	Adm25464 Hypether
43	78	75.7	611	3	AAB33009	Aab33009 Pinus rad
44	78	75.7	611	8	ADS21255	AdS21255 Bacteri
45	77	74.8	648	8	ADS42927	AdS42927 Bacteri

## ALIGNMENTS

RESULT 1	ABW01184	standard; peptide: 23 AA.
ID	ABW01184	standard; peptide: 23 AA.
XX	ABW01184;	
AC	15-JAN-2004.	(first entry)
DT	15-JAN-2004.	(first entry)
XX	Saccharomyces cerevisiae motif VIII peptide.	
DB	Saccharomyces cerevisiae motif VIII peptide.	
XX	Modulator of translation termination; MTT1; helicase B; antiviral;	
KW	therapy; HCSB; nonsense mutation; yeast.	
XX	Saccharomyces cerevisiae.	
OS	Saccharomyces cerevisiae.	
XX	Key	Location/Qualifiers
FT	Misc-difference 4	/label= Unknown
FT	Misc-difference 4	/note= "Xaa may be any amino acid"
FT	Misc-difference 8	/label= Unknown
FT	Misc-difference 8	/note= "Xaa may be any amino acid"
FT	Misc-difference 16	/label= Unknown
FT	Misc-difference 16	/note= "Xaa may be any amino acid"
XX	US6630294-B1.	
XX	07-OCT-2003.	
PD	22-JUL-1999;	99US-00359268.
XX	22-JUL-1999;	98US-0093685P.
PR	22-JUL-1998;	98US-0093685P.
XX	(UTNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.	
PA	Peltz S, Czaplinski K, Dinman JD;	
PI	Peltz S, Czaplinski K, Dinman JD;	
XX	WPI; 2003-810549/76.	
DR	Identifying an agent that increases nonsense suppression, for antiviral	
XX	therapy, by contacting modulator of translation termination (Mtt1) in	
PT	Saccharomyces cerevisiae with a test agent, and detecting specific	
PT	binding to Mtt1.	
XX	Diagnosis; Col 43-44; Opp; English.	
PS	The invention relates to a method of identifying an agent that increases	
XX		
CC		

CC nonsense suppression, by contacting modulator of translation termination  
 CC (MTT1) also referred to as helicase B (HCSB) in *Saccharomyces cerevisiae*.  
 CC The method is useful for identifying compositions or agents which  
 CC increase nonsense suppression. The invention may also be used for  
 CC antiviral therapy and for suppression of pathological nonsense mutations.  
 CC The present sequence is *Saccharomyces cerevisiae* motif peptide  
 CC  
 SQ Sequence 23 AA;

Query Match 94.2%; Score 97; DB 7; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 4e-10; Mismatches 0; Indels 0; Gaps 0;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVTVDXFGGRKDXIIISGVR 23  
 Db 1 VEVTVDXFGGRKDXIIISGVR 23

RESULT 2  
 ADP44110  
 ID ADP44110 standard; peptide; 23 AA.

AC ADP44110;  
 DT 18-NOV-2004 (first entry)

DE Yeast translation termination modulation protein motif VIII.

XX gene therapy; translation termination; RNA helicase; MTT1;  
 XX frameshift frequency; aberrant transcript degradation;  
 KM peptide/l translation; modulation; beta-thalassemia; beta-globin;  
 KM Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
 KM Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;  
 KM Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic fibrosis;  
 KM Kidney Stone; Familial hypercholesterolaemia; Retinitis Pigmentosa;  
 KM Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; Yeast.

XX *Saccharomyces cerevisiae*.  
 OS Synthetic.

PH Key Location/Qualifiers

FT Misc-difference 4 /note= "Any amino acid"

FT Misc-difference 8 /note= "Any amino acid"

FT Misc-difference 16 /note= "Any amino acid"

PN US2004115787-A1.

PD 17-JUN-2004.

PF 28-AUG-2003; 2003US-00652334.

PR 22-JUL-1998; 98US-0093685P.

PR 22-JUL-1999; 99US-00359268.

PA (PELTZ) PELTZ S.

PA (CZAP) CZAPINSKI K.

PA (DINM) DINMAN J D.

PI Pelcz S, Czaplinski K, Dinman JD;

DR WPI; 2004-449400/42.

XX Identifying a test composition or agent that modulates the efficiency of  
 PT translation termination comprises contacting the MTT1 with the test  
 PT composition or agent, and determining if the test composition or agent  
 PT inhibits the MTT1.

PS Claim 40; SEQ ID NO 8; 41pp; English.

XX The invention relates to a method of identifying a test composition that

CC modulates the efficiency of translation termination comprising contacting  
 CC the RNA helicase MTT1 with a composition or agent under conditions  
 CC permitting binding between the MTT1 and the composition, detecting  
 CC specific binding of the test composition or agent to the MTT1, and  
 CC determining if the test composition or agent inhibits the MTT1. The  
 CC composition and methods are useful for modulating the fidelity of  
 CC translation termination or for identifying agents that affect the  
 CC functional activity of mRNAs by altering frameshift frequency, permit  
 CC monitoring of a termination event, promote degradation of aberrant  
 CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl  
 CC transferase activity during initiation, elongation, termination and mRNA  
 CC degradation of translation. The agents, which may be antagonists or  
 CC agonists, are useful in screening, diagnostic and therapeutic purposes,  
 CC for diseases or conditions resulting from or cause premature translation,  
 CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular  
 CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,  
 CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,  
 CC Hirschsprung disease, Cystic fibrosis, Kidney Stones, Familial  
 CC hypercholesterolaemia, Retinitis Pigmentosa, or Neurofibromatosis,  
 CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents  
 CC the amino acid sequence of the yeast translation termination modulation  
 CC protein motif VIII.

XX Sequence 23 AA;

Query Match 94.2%; Score 97; DB 8; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 4e-10; Mismatches 0; Indels 0; Gaps 0;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVTVDXFGGRKDXIIISGVR 23  
 Db 1 VEVTVDXFGGRKDXIIISGVR 23

RESULT 3  
 AAY77811  
 ID AAY77811 standard; peptide; 24 AA.

AC AAY77811;

DT 31-MAY-2000 (first entry)

DE Motif VIII comprised in a gene modulating translation termination.

XX Helicase B; HCSB; MTT1; modulator of translation termination; eRF1; eRF3;  
 KM eukaryotic release factor; peptidyl transferase; beta-thalassemia;  
 KM beta-globin; Duchenne/Becker Muscular Dystrophy; antianemic.

XX Unidentified.

OS Unidentified.

PH Key Location/Qualifiers

FT Misc-difference 1.24 /note= "residues indicated Xaa are unspecified"

PN WO200005586-A2.

PD 03-FEB-2000.

PR 22-JUL-1999; 99NO-US016802.

PR 22-JUL-1998; 98US-00120435.

PA (UYNE-) UNIV NEW JERSEY.

PI Pelcz S, Czaplinski K, Dinman JD;

DR WPI; 2000-171458/15.

XX New multiprotein complex which can modulate peptidyl transferase activity  
 PT during translation, useful to treat diseases associated with peptidyl  
 PT transferase activity e.g. Duchenne/Becker Muscular Dystrophy.

PS Claim 40; Page 80; 89pp; English.



XX The invention provides a new multiprotein complex which can modulate  
 CC peptidyl transferase activity during translation. The complex comprises  
 CC the gene encoding Helicase B (HCSB; renamed MTT1, for Modulator of  
 CC Translation Termination) and the conserved proteins known to interact and  
 CC carry out translation termination in eukaryotic cells, peptidyl  
 CC eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to  
 CC modulate peptidyl transferase activity during translation in a cell. It  
 CC can be administered therapeutically combined with a carrier in  
 CC pharmaceutical compositions to treat diseases associated with peptidyl  
 CC transferase activity, especially diseases resulting from a nonsense or  
 CC frameshift mutation e.g. beta-thalassemia, beta-globin, Duchenne/Becker  
 CC Muscular Dystrophy etc. It can be used to identify disease conditions  
 CC involving a defect in the complex, by transfecting cells with encoding  
 CC nucleic acid and determining the proportion of defective complex before  
 CC and after transfection. It is also useful to screen for drugs involved in  
 CC peptidyl transferase activity during translation, inhibiting the  
 CC interaction between MTT1 and eRF3 or involved in enhancing translation  
 CC termination. Vectors comprising polynucleotides encoding the complex (or  
 CC antisense sequences) can be constructed and introduced into cells to  
 CC interfere with complex expression and so modulate the efficiency of  
 CC translation termination of mRNA and/or degradation of aberrant  
 CC transcripts in a cell. Agents binding to the complex can be identified  
 CC and included in therapeutic compositions useful as above, and/or used to  
 CC modulate peptidyl transferase activity during translation in cells. They  
 CC are also useful to modulate the efficiency of translation termination of  
 CC mRNA at a nonsense codon and/or promote degradation of aberrant  
 CC transcripts in cells. The method can be used to identify agents/  
 CC compositions modulating binding to MTT1, useful to identify genes.  
 CC Sequences AA177804-812 represent motifs I-IX comprised in the genes of  
 CC interest, used for modulating translation termination

SO Sequence 24 AA;

Query Match 94.2%; Score 97; DB 3; Length 24;

Best Local Similarity 100.0%; Pred. No. 4.2e-10;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VEVXTVDXFGGREKDXITLSCVR 23  
 Db 1 VEVXTVDXFGGREKDXITLSCVR 23

RESULT 4

AA177814  
 ID AA177814 standard; peptide; 380 AA.

XX AA177814;

DT 31-MAY-2000 (first entry)

DE Yeast Upfl protein fragment.

XX Helicase B; HCSB; MTT1; modulator of translation termination; eRF1, eRF3;

KW eukaryotic release factor; peptidyl transferase; beta-thalassemia;

KW beta-globin; Duchenne/Becker Muscular Dystrophy; antinematic; yeast;

KW helicase; Upfl.

XX Saccharomyces cerevisiae.

PN WO200005586-A2.

PD 03-FEB-2000.

PP 22-JUL-1999; 99WO-US016802.

PR 22-JUL-1998; 98US-00120435.

PA (UYNE-) UNIV NEW JERSEY.

PI Peletz S, Czaplinski K, Dinman JD;

DR WPI; 2000-171458/15.

XX New multiprotein complex which can modulate peptidyl transferase activity  
 PT during translation, useful to treat diseases associated with peptidyl  
 PT transferase activity e.g. Duchenne/Becker Muscular Dystrophy.

PS Example 1; Fig 1; 89pp; English.

XX The invention provides a new multiprotein complex which can modulate  
 CC peptidyl transferase activity during translation. The complex comprises  
 CC the gene encoding Helicase B (HCSB; renamed MTT1, for Modulator of  
 CC Translation Termination) and the conserved proteins known to interact and  
 CC carry out translation termination in eukaryotic cells, peptidyl  
 CC eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to  
 CC modulate peptidyl transferase activity during translation in a cell. It  
 CC can be administered therapeutically combined with a carrier in  
 CC pharmaceutical compositions to treat diseases associated with peptidyl  
 CC transferase activity, especially diseases resulting from a nonsense or  
 CC frameshift mutation e.g. beta-thalassemia, beta-globin, Duchenne/Becker  
 CC Muscular Dystrophy etc. It can be used to identify disease conditions  
 CC involving a defect in the complex, by transfecting cells with encoding  
 CC nucleic acid and determining the proportion of defective complex before  
 CC and after transfection. It is also useful to screen for drugs involved in  
 CC peptidyl transferase activity during translation, inhibiting the  
 CC interaction between MTT1 and eRF3 or involved in enhancing translation  
 CC termination. Vectors comprising polynucleotides encoding the complex (or  
 CC antisense sequences) can be constructed and introduced into cells to  
 CC interfere with complex expression and so modulate the efficiency of  
 CC translation termination of mRNA and/or degradation of aberrant  
 CC transcripts in a cell. Agents binding to the complex can be identified  
 CC and included in therapeutic compositions useful as above, and/or used to  
 CC modulate peptidyl transferase activity during translation in cells. They  
 CC are also useful to modulate the efficiency of translation termination of  
 CC mRNA at a nonsense codon and/or promote degradation of aberrant  
 CC transcripts in cells. The method can be used to identify agents/  
 CC compositions modulating binding to MTT1, useful to identify genes.  
 CC Sequences AA177813-817 represent protein fragments from yeast superfamily  
 CC group I helicases

SO Sequence 380 AA;

Query Match 92.2%; Score 95; DB 3; Length 380;

Best Local Similarity 82.6%; Pred. No. 2.2e-08;

Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 VEVXTVDXFGGREKDXITLSCVR 23  
 Db 334 VEVASVDXFGGREKDXITLSCVR 356

RESULT 5

ABW01205  
 ID ABW01205 standard; protein; 380 AA.

XX ABW01205;

DT 15-JAN-2004 (first entry)

DE Saccharomyces cerevisiae UPfl protein.

XX Modulator of translation termination; MTT1; helicase B; antiviral;

KW therapy; HCSB; nonsense mutation; yeast; RENT1; HUPF1; UPfl.

XX Saccharomyces cerevisiae.

PN US6630294-B1.

PD 07-OCT-2003.

PP 22-JUL-1999; 99US-00359268.

PR 22-JUL-1998; 98US-0093685P.

PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

```

XX PI Peltz S, Czaplinski K, Dimman JD;
XX WPI; 2003-810549/76.
XX
XX PT Identifying an agent that increases nonsense suppression, for antiviral
XX therapy, by contacting modulator of translation termination (MTT1) in
XX Saccharomyces cerevisiae with a test agent, and detecting specific
XX binding to MTT1.
XX PS Disclosure; Col 57-60; Opp; English.
XX
XX CC The invention relates to a method of identifying an agent that increases
XX nonsense suppression, by contacting modulator of translation termination
XX (MTT1) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.
XX The method is useful for identifying compositions or agents which
XX increase nonsense suppression. The invention may also be used for
XX antiviral therapy and for suppression of pathological nonsense mutations.
XX The present sequence is Saccharomyces cerevisiae UPF1 (otherwise termed
XX RENT1, HUPF1) protein
XX
XX SQ Sequence 380 AA;
XX
XX Query Match 92.2%; Score 95; DB 7; Length 380;
XX Best Local Similarity 82.6%; Pred. No. 2.2e-08;
XX Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 1 VEVYTVDXFGQREKDXITLLSCVR 23
XX |||:|||||
XX DB 334 VEVASVDAFGQREKDXITLLSCVR 356
XX
XX RESULT 6
XX ID ADP44131 standard; protein; 380 AA.
XX AC ADP44131;
XX
XX DT 18-NOV-2004 (first entry)
XX
XX DE Yeast helicase Upf1.
XX
XX KW gene therapy; translation termination; RNA helicase; MTT1;
XX frameshift frequency; aberrant transcript degradation;
XX peptidyl transferase modulation; beta-thalassemia; beta-globin;
XX Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;
XX Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;
XX Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic fibrosis;
XX Kidney Stone; Familial hypercholesterolaemia; Retinitis pigmentosa;
XX Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; yeast; enzyme.
XX
XX OS Saccharomyces cerevisiae.
XX
XX PN US204115787-A1.
XX
XX PD 17-JUN-2004.
XX
XX PF 28-AUG-2003; 2003US-00652334.
XX
XX PR 22-JUL-1998; 98US-0093685P.
XX
XX PR 22-JUL-1999; 99US-00359268.
XX
XX PA (PELTZ/) PELTZ S.
XX PA (CZAP/) CZAPINSKI K.
XX PA (DINM/) DINMAN J D.
XX
XX PI Peltz S, Czaplinski K, Dimman JD;
XX WPI; 2004-449400/42.
XX
XX PT Identifying a test composition or agent that modulates the efficiency of
XX translation termination comprises contacting the MTT1 with the test
XX composition or agent, and determining if the test composition or agent

```

```

PT inhibits the MTT1.
XX
XX PS Disclosure; SEQ ID NO 29; 41pp; English.
XX
XX CC The invention relates to a method of identifying a test composition that
XX modulates the efficiency of translation termination comprising contacting
XX the RNA helicase MTT1 with a composition or agent under conditions
XX permitting binding between the MTT1 and the composition, detecting
XX specific binding of the test composition or agent to the MTT1, and
XX determining if the test composition or agent inhibits the MTT1. The
XX composition and methods are useful for modulating the fidelity of
XX translation termination or for identifying agents that affect the
XX functional activity of mRNAs by altering frameshift frequency, permit
XX monitoring of a termination event, promote degradation of aberrant
XX transcripts, and provide modulators (inhibitors/activators) of peptidyl
XX transferase activity during initiation, elongation, termination and mRNA
XX degradation of translation. The agents, which may be antagonists or
XX agonists, are useful in screening, diagnostic and therapeutic purposes,
XX for diseases or conditions resulting from or cause premature translation,
XX such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular
XX Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,
XX Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,
XX Hirschsprung disease, Cystic fibrosis, Kidney Stones, Familial
XX hypercholesterolaemia, Retinitis pigmentosa, or Neurofibromatosis,
XX Retinoblastoma, ATM or Costmann Disease. The present sequence represents
XX the amino acid sequence of the yeast helicase Upf1.
XX
XX SQ Sequence 380 AA;
XX
XX Query Match 92.2%; Score 95; DB 8; Length 380;
XX Best Local Similarity 82.6%; Pred. No. 2.2e-08;
XX Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 1 VEVYTVDXFGQREKDXITLLSCVR 23
XX |||:|||||
XX DB 334 VEVASVDAFGQREKDXITLLSCVR 356
XX
XX RESULT 7
XX ID ADN19434 standard; protein; 925 AA.
XX AC ADN19434;
XX
XX DT 02-DEC-2004 (first entry)
XX
XX DB Bacterial polypeptide #2087.
XX
XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polypeptide.
XX
XX OS Bacteria.
XX
XX PN US2003233675-A1.
XX
XX PD 18-DEC-2003.
XX
XX PF 20-FEB-2003; 2003US-00369493.
XX
XX PR 21-FEB-2002; 2002US-0360039P.
XX
XX PA (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.
XX
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

```

XX WP1; 2004-061375/06.  
 XX  
 PT New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.  
 XX  
 PS Claim 1; SEQ ID NO 2087; 122pp; English.  
 XX  
 CC The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition. Improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at seqdata.uspto.gov/sequence.html.  
 CC  
 XX  
 SQ Sequence 925 AA;  
 Query Match 92.2%; Score 95; DB 8; Length 925;  
 Best Local Similarity 82.6%; Pred. No. 5.9e-08;  
 Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 VEVXTVDXFOGREKDXIIILSCVR 23  
 Db 749 VEVASVDXFOGREKDXIIILSCVR 771  
 RESULT 8  
 AAY98057  
 ID AAY98057 standard; protein; 971 AA.  
 XX  
 AC AAY98057;  
 XX  
 DT 07-SEP-2000 (first entry)  
 XX  
 DE Yeast Upf1p.  
 XX  
 XX Yeast; nonsense-mediated mRNA decay pathway; NMD2; UPF3; UPF1; upf1p;  
 KM upf1p; gene therapy; beta-thalassemia; cancer; Fanconi anaemia;  
 KM polycystic kidney disease; Duchenne muscular dystrophy; Tay-Sachs disease;  
 KM hemophilia; hypercholesterolemia; neurofibromatosis; Tay-Sachs disease;  
 KM glycogen storage disease; cystic fibrosis; adenomatous polyposis coli;  
 KM Cowden disease; Maple syrup urine disease; Wilson disease;  
 KM Niemann-Pick disease; Turcot syndrome; McArdle disease;  
 KM ornithine transcarbamylase deficiency.  
 KM  
 OS Saccharomyces cerevisiae.  
 OS  
 XX  
 PN US6071700-A.  
 XX  
 PD 06-JUN-2000.  
 XX  
 XX 21-OCT-1998; 98US-00177431.  
 XX  
 PR 20-JAN-1995; 95US-00375300.  
 PR

PR 21-OCT-1997; 97US-00955472.  
 XX  
 PA (UYMA-) UNIV MASSACHUSETTS.  
 XX  
 PI Jacobson AS, He P;  
 XX  
 DR WP1; 2000-422078/36.  
 DR N-PSDB; AAA39454.  
 XX  
 XX Testing for compounds able to modulate the nonsense mediated mRNA decay  
 PT pathway for increasing endogenous protein production by incubating a  
 PT candidate compound with a cell having a gene with a nonsense mutation and  
 PT a gene without mutation.  
 XX  
 PS Disclosure; Fig 6; 57pp; English.  
 XX  
 CC The nonsense-mediated mRNA decay pathway has an important role in the  
 CC degradation of mRNA transcripts that contain a stop codon. The products  
 CC of the UPF1 and UPF3 genes (protein Upf1p and Upf3p, respectively) are  
 CC essential components of this pathway. Mutations in these genes or  
 CC inhibition of Upf1p or Upf3p function would lead to stabilisation of mRNA  
 CC containing premature stop codons. The present sequence is the UPF1  
 CC protein, Upf1p, from Saccharomyces cerevisiae. The carboxy terminal of  
 CC the NMD2 gene protein, Nmd2p, has been found to bind to Upf1p, leading to  
 CC inhibition of the nonsense-mediated mRNA decay pathway. Inhibition of  
 CC this pathway is a useful means of treating disorders caused by the  
 CC presence of nonsense mutations, e.g. breast cancer, polycystic kidney  
 CC diseases, Niemann-Pick disease, adenomatous polyposis coli, cystic  
 CC fibrosis, Fanconi's anaemia, haemophilia, hypercholesterolemia,  
 CC neurofibromatosis, ornithine transcarbamylase deficiency, retinoblastoma,  
 CC glycogen storage disease, McArdle disease, colorectal cancer, Tay-Sachs  
 CC disease, Cowden disease, Wilson disease, polycystic kidney disease,  
 CC Duchenne muscular dystrophy, adrenocortical carcinoma, Li-Fraumeni  
 CC syndrome, Maple syrup urine disease, Turcot syndrome or beta-thalassemia  
 CC  
 XX  
 SQ Sequence 971 AA;  
 Query Match 92.2%; Score 95; DB 3; Length 971;  
 Best Local Similarity 82.6%; Pred. No. 6.3e-08;  
 Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 VEVXTVDXFOGREKDXIIILSCVR 23  
 Db 757 VEVASVDXFOGREKDXIIILSCVR 779  
 RESULT 9  
 ABR53412  
 ID ABR53412 standard; protein; 971 AA.  
 XX  
 AC ABR53412;  
 XX  
 DT 20-JUN-2003 (first entry)  
 XX  
 DE Protein sequence #SEQ ID 1689.  
 XX  
 XX Multiprotein complex; eukaryote; drug target; diagnosis.  
 KM  
 KM Saccharomyces cerevisiae.  
 OS  
 OS EPI258494-A1.  
 PN  
 PD 20-NOV-2002.  
 XX  
 XX 20-DEC-2001; 2001EP-00130253.  
 XX  
 XX 15-MAY-2001; 2001BP-00111774.  
 XX  
 XX (CELL-) CELZOME AG.  
 XX  
 XX Bauer A, Gavin A, Grandt P, Krause R, Kruse UD, Kuester BD;  
 PI Marzloch M, Schultz JD, Superti-Furga GD;  
 XX

DR WPI; 2003-25078/25.  
DR N-PSDB; ACC61454.  
XX  
PT New isolated protein complexes useful for diagnosing a disease or  
PT disorder, or as a target for an active agent of a pharmaceutical,  
PT preferably a drug target in the treatment or prevention of disease or  
PT disorder.  
XX  
PS Disclosure; SEQ ID NO 1689; 17pp + Sequence Listing; English.  
XX  
CC The invention relates to multiprotein complexes from eukaryotes. Proteins  
CC of the invention and DNA sequences encoding them are given in records  
CC ABR5568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are  
CC obtainable by using a protein as a bait and isolating the set of proteins  
CC which is attached thereto from cells. Such protein complexes may comprise  
CC up to 30 distinct proteins. Protein complexes of the invention are useful  
CC for diagnosing a disease or disorder, or as a target for an active agent  
CC of a pharmaceutical, preferably a drug target in the treatment or  
CC prevention of a disease or disorder. Note: The sequence data for this  
CC patent is not represented in the printed specification, but is based on  
CC sequence information supplied by the European Patent Office. The complete  
CC document is available on CD-ROM  
XX  
SQ Sequence 971 AA;  
XX  
Query Match 92.2%; Score 95; DB 6; Length 971;  
Best Local Similarity 82.6%; Pred. No. 6,3e-08;  
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
XX  
QY 1 VEVXYVDXFGGREKDXIIILSCVR 23  
ID ADR10398 standard; protein; 373 AA.  
XX  
AC ADR10398;  
XX  
DT 04-NOV-2004 (first entry)  
XX  
DE Human protein useful for treating neurological disease Seq 3904.  
XX  
KM human; oligo-capping method; diagnostic marker; gene therapy;  
KM osteoporosis; neurological disease; Alzheimer's disease;  
KM Parkinson's disease; dementia; short memory; cancer;  
KM sense or motor function; emotional reaction; fear response; panic;  
KM osteopathic; neuroprotective; nootropic; antiparkinsonian; cyostatic;  
KM tranquiliser.  
XX  
OS Homo sapiens.  
XX  
PN EP1447413-A2.  
XX  
PD 18-AUG-2004.  
XX  
PF 12-FEB-2004; 2004EP-00003145.  
XX  
PR 14-FEB-2003; 2003JP-00102207.  
XX  
PR 09-MAY-2003; 2003JP-00131452.  
XX  
PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;  
PI Wakamatsu A, Ishii S, Nagai K, Irie R;  
PI Wakamatsu A, Ishii S, Nagai K, Irie R;  
XX  
DR WPI; 2004-583265/57.  
DR N-PSDB; ADR08442.  
XX  
XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,  
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.  
XX  
PS Claim 1; SEQ ID NO 3904; 2686pp; English.  
XX  
CC This invention relates to novel, isolated full length human cDNA  
CC molecules and the encoded proteins thereof. Specifically, it refers to  
CC cDNA clones obtained by an oligo-capping method, where none of these

CC second protein, or its derivative, fragment, homologue or variant. The  
CC proteins are selected from given protein complexes, which are not defined  
CC in the specification. The variants are encoded by nucleic acids that  
CC hybridize to the nucleic acids encoding the proteins under low stringency  
CC conditions. The protein complexes are useful as targets for an active  
CC agent of a pharmaceutical. These protein complexes are particularly  
CC useful as drug targets for the treatment or prevention of a disease or  
CC disorder. The complexes and methods above are useful in diagnosing or  
CC screening for the presence of a disease or disorder or a predisposition  
CC for developing a disease or disorder in a subject. These are also useful  
CC in screening for a drug for treatment or prevention of a disease or  
CC disorder. The molecule that modulates the amount, activity or protein  
CC components of the complex is useful for the manufacture of a medicament  
CC for the treatment or prevention of a disease or disorder. This sequence  
CC corresponds to a protein of the invention. (Note: the sequence data for  
CC this patent did not form part of the printed specification but was  
CC obtained from the EPO in electronic format).  
XX  
SQ Sequence 971 AA;  
XX  
Query Match 92.2%; Score 95; DB 7; Length 971;  
Best Local Similarity 82.6%; Pred. No. 6,3e-08;  
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
XX  
QY 1 VEVXYVDXFGGREKDXIIILSCVR 23  
ID ADR10398 standard; protein; 373 AA.  
XX  
AC ADR10398;  
XX  
DT 04-NOV-2004 (first entry)  
XX  
DE Human protein useful for treating neurological disease Seq 3904.  
XX  
KM human; oligo-capping method; diagnostic marker; gene therapy;  
KM osteoporosis; neurological disease; Alzheimer's disease;  
KM Parkinson's disease; dementia; short memory; cancer;  
KM sense or motor function; emotional reaction; fear response; panic;  
KM osteopathic; neuroprotective; nootropic; antiparkinsonian; cyostatic;  
KM tranquiliser.  
XX  
OS Homo sapiens.  
XX  
PN EP1447413-A2.  
XX  
PD 18-AUG-2004.  
XX  
PF 12-FEB-2004; 2004EP-00003145.  
XX  
PR 14-FEB-2003; 2003JP-00102207.  
XX  
PR 09-MAY-2003; 2003JP-00131452.  
XX  
PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;  
PI Wakamatsu A, Ishii S, Nagai K, Irie R;  
PI Wakamatsu A, Ishii S, Nagai K, Irie R;  
XX  
DR WPI; 2004-583265/57.  
DR N-PSDB; ADR08442.  
XX  
XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,  
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.  
XX  
PS Claim 1; SEQ ID NO 3904; 2686pp; English.  
XX  
CC This invention relates to novel, isolated full length human cDNA  
CC molecules and the encoded proteins thereof. Specifically, it refers to  
CC cDNA clones obtained by an oligo-capping method, where none of these

CC clones are identical to any known human mRNAs. The present invention  
 CC describes an immunoassay to identify agonists and antagonists, as well as  
 CC antibodies, antisense molecules and siRNAs that can all be used to bind  
 CC to and modulate expression of the cDNA molecules. As such, these  
 CC molecules are useful for diagnostic markers or therapeutic targets for  
 CC the various diseases or morbid states. In particular, they are useful in  
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's  
 CC disease, Parkinson's disease, dementia, short memory and various cancers,  
 CC as well as for maintaining equilibrium of sense or motor function, and  
 CC for treating emotional reaction, fear response and panic. Accordingly,  
 CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,  
 CC cyostatic and tranquilliser activities. This polypeptide is a protein  
 CC encoded by a full length human cDNA sequence of the invention. NOTE: This  
 CC sequence is not given in the sequence listing of the specification but  
 CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-  
 CC office.

XX Sequence 373 AA;

Query Match 91.3%; Score 94; DB 8; Length 373;  
 Best Local Similarity 78.3%; Pred. No. 3.2e-08;  
 Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VEVYTVDXFGGRKDXIIISLCVR 23  
 ||:|||||  
 Db 76 VEIASVDAFGGRKDXIIISLCVR 98

RESULT 12

AAW36508 standard; protein; 797 AA.

AC AAW36508;

DT 14-JUN-1998 (first entry)

DE Human RENT1 protein fragment #5.

KM RENT1, nonsense-mediated RNA decay; NMRD; mutation; diagnosis; therapy;  
 KW regulator of nonsense transcripts; Marfan Syndrome; aging; cancer.

XX Homo sapiens.

PN MO9740855-A1.

PD 06-NOV-1997.

PF 01-OCT-1996; 96WO-US015769.

PT 29-APR-1996; 96US-0016482P.

PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.

PI Dietz HC;

DR WPI; 1997-549494/50.

PT New isolated regulators of non-sense-mediated RNA decay - used to develop  
 PT products for the study, diagnosis and therapy of disorders such as Marfan  
 PT Syndrome, accelerated aging and cancers.

PS Claim 1; Fig 1B; 79pp; English.

CC This protein sequence is a fragment of the human RENT1 protein which  
 CC regulates nonsense-mediated RNA decay (NMRD). This fragment of RENT1  
 CC shows homology to the yeast Upf1 protein. The RENT1 (regulator of  
 CC nonsense transcripts) protein and other products can be used in the  
 CC study, diagnosis and therapy of disorders involving NMRD such as Marfan  
 CC Syndrome, accelerated aging or various cancers

XX Sequence 797 AA;

Query Match 91.3%; Score 94; DB 2; Length 797;

Best Local Similarity 78.3%; Pred. No. 7.6e-08;  
 Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VEVYTVDXFGGRKDXIIISLCVR 23  
 ||:|||||  
 Db 701 VEIASVDAFGGRKDXIIISLCVR 723

RESULT 13

ADJ69900 standard; protein; 935 AA.

AC ADJ69900;

DT 06-MAY-2004 (first entry)

DE Human heart mitochondrial protein as a therapeutic target SegID1706.

XX mitochondrial; human; screening assay; diabetes mellitus;

KM Huntington's disease; osteoarthritis;

KM Leber's hereditary optic neuropathy; LHON;  
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;  
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
 KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;  
 KW osteopathic; ophthalmological; cytostatic.

XX

OS Homo sapiens.

PN WO2003087768-A2.

PD 23-OCT-2003.

PE 04-APR-2003; 2003WO-US010870.

PR 12-APR-2002; 2002US-0372843P.

PR 17-JUN-2002; 2002US-0389987P.

PR 20-SEP-2002; 2002US-0412418P.

PA (MITO-) MITOKOR.

PI (BUCK-) BUCK INST AGE RES.

PI Warnock DB;

DR WPI; 2003-845369/78.

PT Identifying a mitochondrial target for drug screening assays and for  
 PT treating diseases associated with altered mitochondrial function,  
 PT comprises detecting a modified polypeptide in a sample and correlating  
 PT with the disease.

PS Claim 1; SEQ ID NO 1706; 180pp; English.

CC This invention relates to novel mitochondrial targets that can be used  
 CC for therapeutic intervention in treating a disease associated with  
 CC altered mitochondrial function. Specifically, it refers to a method for  
 CC identifying proteins of the human heart mitochondrial proteome that are  
 CC useful for drug screening assays, as well as therapeutic targets. The  
 CC present invention describes a method for identifying such proteins that  
 CC can be used in the treatment of various diseases associated with altered  
 CC mitochondrial function including diabetes mellitus, Huntington's disease,  
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
 CC compositions have neuroprotective, nootropic, antidiabetic,  
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
 CC cyostatic activities. This polypeptide sequence is a human heart  
 CC mitochondrial protein of the invention.

XX Sequence 935 AA;

Query Match 91.3%; Score 94; DB 7; Length 935;  
 Best Local Similarity 78.3%; Pred. No. 9.1e-08;



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 07:37:21 ; Search time 28.708 Seconds  
(without alignments)  
59.807 Million cell updates/sec

Title: US-10-652-334-8

Sequence: 1 VEVXTVDXFOGKREKDXIIISCVR 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	94.2	23	4	US-09-359-268A-8
2	95	92.2	380	4	US-09-359-268A-29
3	95	92.2	971	2	US-08-724-354D-22
4	95	92.2	971	3	US-09-270-984A-22
5	95	92.2	971	3	US-09-177-421-8
6	94	91.3	1043	2	US-08-724-354D-4
7	94	91.3	1043	3	US-09-270-984A-4
8	94	91.3	1118	2	US-08-724-354D-2
9	94	91.3	1118	3	US-09-270-984A-2
10	94	91.3	1140	4	US-09-949-016-10116
11	86	83.5	414	4	US-09-359-268A-28
12	86	83.5	683	4	US-09-538-092-483
13	86	79.6	211	4	US-09-902-540-16148
14	80	77.7	917	4	US-09-248-796A-19347
15	79	76.7	426	4	US-09-248-796A-15170
16	78	75.7	171	4	US-09-640-211A-1058
17	78	75.7	405	4	US-09-248-796A-19107
18	76	73.8	415	4	US-09-359-268A-25
19	75	72.8	472	4	US-09-359-268A-26
20	74	71.8	993	4	US-09-538-092-1100
21	66	64.1	219	4	US-09-248-796A-18993
22	66	64.1	366	4	US-09-359-268A-27
23	61	59.2	157	4	US-09-270-767-32463
24	61	59.2	157	4	US-09-270-767-47680
25	61	59.2	3177	2	US-08-477-451-4
26	60	58.3	181	4	US-09-270-767-31838
27	60	58.3	181	4	US-09-270-767-47055

28	56	54.4	85	4	US-09-270-767-59862	Sequence 59862, A
29	56	54.4	444	4	US-09-270-767-44429	Sequence 44429, A
30	53	51.5	11	4	US-09-359-268A-21	Sequence 21, Appl
31	49	47.6	1020	4	US-09-328-352-6682	Sequence 6682, Ap
32	42	40.8	1345	4	US-09-949-016-8313	Sequence 8313, Ap
33	42	40.8	2213	1	US-08-727-034-3	Sequence 3, Appl
34	42	40.8	2214	1	US-08-727-034-7	Sequence 40, Appl
35	42	40.8	2214	4	US-09-919-039-40	Sequence 37, Appl
36	40	38.8	58	4	US-09-549-441-37	Sequence 8648, Ap
37	40	38.8	142	4	US-09-489-039A-8648	Sequence 4, Appl
38	40	38.8	497	3	US-09-058-947A-4	Sequence 12, Appl
39	40	38.8	500	3	US-08-868-373-12	Sequence 2, Appl
40	40	38.8	1833	3	US-08-479-722B-2	Sequence 7, Appl
41	40	38.8	1833	4	US-09-592-685-2	Sequence 18, Appl
42	40	38.8	1833	5	PCT-US95-02251-18	Sequence 3685, Ap
43	39	37.9	226	1	US-09-134-000C-3685	Sequence 14, Appl
44	39	37.9	307	1	US-08-597-226-14	Sequence 14, Appl
45	39	37.9	307	1	US-08-746-682A-14	

#### ALIGNMENTS

```
RESULT 1
US-09-359-268A-8.
Sequence 8, Application US/09359268A
Patent No. 6630294
GENERAL INFORMATION:
APPLICANT: Peltz, Stuart
APPLICANT: Czaplinski, Kevin
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
FILE REFERENCE: 601-1-85N
CURRENT FILING DATE: 1999-07-22
PRIOR APPLICATION NUMBER: 60/093,685
PRIOR FILING DATE: 1998-07-22
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 8
LENGTH: 23
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
FEATURES:
OTHER INFORMATION: Xaa = any amino acid
US-09-359-268A-8
Query Match 94.2%; Score 97; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 7.6e-12; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;
Cy 1 VEVXTVDXFOGKREKDXIIISCVR 23
Db 1 VEVXTVDXFOGKREKDXIIISCVR 23
RESULT 2
US-09-359-268A-29
Sequence 29, Application US/09359268A
Patent No. 6630294
GENERAL INFORMATION:
APPLICANT: Peltz, Stuart
APPLICANT: Czaplinski, Kevin
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
FILE REFERENCE: 601-1-85N
CURRENT FILING DATE: 1999-07-22
CURRENT APPLICATION NUMBER: US/09/359,268A
PRIOR APPLICATION NUMBER: 60/093,685
```



PRIOR FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 29  
LENGTH: 380  
TYPE: PRT  
ORGANISM: saccharomyces cerevisiae  
US-09-359-268A-29

Query Match 92.2%; Score 95; DB 4; Length 380;  
Best Local Similarity 82.6%; Pred. No. 5.2e-10;  
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEVXTVDXFOGREKDXIILSCVR 23  
|||:|||||  
Db 334 VEVASVDAFOGREKDXIILSCVR 356

## RESULT 3

US-08-724-354D-22  
Sequence 22, Application US/08724354D  
Patent No. 5994119  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Excutive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,354D  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/016,482  
FILING DATE: 29-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 971 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-724-354D-22

Query Match 92.2%; Score 95; DB 2; Length 971;  
Best Local Similarity 82.6%; Pred. No. 1.6e-09;  
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEVXTVDXFOGREKDXIILSCVR 23  
|||:|||||  
Db 757 VEVASVDAFOGREKDXIILSCVR 779

## RESULT 4

US-09-270-984A-22  
Sequence 22, Application US/09270984A  
Patent No. 6048965

GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Excutive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/270,984A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/724,354  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 971 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-270-984A-22

Query Match 92.2%; Score 95; DB 3; Length 971;  
Best Local Similarity 82.6%; Pred. No. 1.6e-09;  
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEVXTVDXFOGREKDXIILSCVR 23  
|||:|||||  
Db 757 VEVASVDAFOGREKDXIILSCVR 779

## RESULT 5

US-09-177-431-8  
Sequence 8, Application US/09177431  
Patent No. 6071700  
GENERAL INFORMATION:  
APPLICANT: He, Feng  
APPLICANT: Jacobson, Allan S.  
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE  
TITLE OF INVENTION: ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/177,431  
FILING DATE:  
PRIOR APPLICATION DATA:



```

: APPLICATION NUMBER: 08/955,472
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Paase, J. Peter
: REGISTRATION NUMBER: 32,983
: REFERENCE/DOCKET NUMBER: 07917/050001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617/542-5070
: TELEFAX: 617/542-9806
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 971 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FRAGMENT TYPE: internal
: US-09-177-431-8

Query Match          92.2%; Score 95; DB 3; Length 971;
Best Local Similarity 82.6%; Pred. No. 1.6e-09;
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 VEVXTVDXFOGREKDXIIILSCVR 23
      ||:|||||
Db      757 VEVASVDAFOGREKDXIIILSCVR 779

RESULT 6
US-08-724-354D-4
: Sequence 4, Application US/08724354D
: Patent No. 5994119
: GENERAL INFORMATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Dietz, Harry C.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 07265/090001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-678-5099
: TELEFAX: 619-678-5099
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1043 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-724-354D-4

Query Match          91.3%; Score 94; DB 2; Length 1043;
Best Local Similarity 78.3%; Pred. No. 2.7e-09;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```

QY      1 VEVXTVDXFOGREKDXIIILSCVR 23
      ||:|||||
Db      744 VEIASVDAFOGREKDXIIILSCVR 766

RESULT 7
US-09-270-984A-4
: Sequence 4, Application US/09270984A
: Patent No. 6048965
: GENERAL INFORMATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Dietz, Harry C.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 07265/090001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-678-5099
: TELEFAX: 619-678-5099
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1043 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-270-984A-4

Query Match          91.3%; Score 94; DB 3; Length 1043;
Best Local Similarity 78.3%; Pred. No. 2.7e-09;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 VEVXTVDXFOGREKDXIIILSCVR 23
      ||:|||||
Db      744 VEIASVDAFOGREKDXIIILSCVR 766

RESULT 8
US-08-724-354D-2
: Sequence 2, Application US/08724354D
: Patent No. 5994119
: GENERAL INFORMATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Dietz, Harry C.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 07265/090001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-678-5099
: TELEFAX: 619-678-5099
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1043 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-724-354D-2

Query Match          91.3%; Score 94; DB 3; Length 1043;
Best Local Similarity 78.3%; Pred. No. 2.7e-09;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,354D
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/016,482
FILING DATE: 29-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1118 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-724-354D-2

Query Match      91.3%; Score 94; DB 2; Length 1118;
Best Local Similarity 78.3%; Pred. No. 2.9e-09;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 VEVXVDXFGQREKDXIIISCVR 23
Db      821 VEIASVDAPQGREDFITLSCVR 843

RESULT 9
US-09-270-984A-2
Sequence 2, Application US/09270984A
Patent No. 6048965
GENERAL INFORMATION:
APPLICANT: Dietz, Harry C.
TITLE OF INVENTION: MAMMALIAN REGULATOR OF
TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/270,984A
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/724,354
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1118 amino acids
TYPE: amino acid
```

```
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-270-984A-2

Query Match      91.3%; Score 94; DB 3; Length 1118;
Best Local Similarity 78.3%; Pred. No. 2.9e-09;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 VEVXVDXFGQREKDXIIISCVR 23
Db      821 VEIASVDAPQGREDFITLSCVR 843

RESULT 10
US-09-949-016-10116
Sequence 10116, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10116
LENGTH: 1140
TYPE: PRT
ORGANISM: Human
US-09-949-016-10116

Query Match      91.3%; Score 94; DB 4; Length 1140;
Best Local Similarity 78.3%; Pred. No. 3e-09;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 VEVXVDXFGQREKDXIIISCVR 23
Db      843 VEIASVDAPQGREDFITLSCVR 865

RESULT 11
US-09-359-268A-28
Sequence 28, Application US/09359268A
Patent No. 6630294
GENERAL INFORMATION:
APPLICANT: Peltz, Stuart
APPLICANT: Czaplinski, Kevin
APPLICANT: Dimman, Jonathan D.
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
FILE REFERENCE: 601-1-85N
CURRENT APPLICATION NUMBER: US/09/359,268A
CURRENT FILING DATE: 1999-07-22
PRIOR APPLICATION NUMBER: 60/093,685
PRIOR FILING DATE: 1998-07-22
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.0.
SEQ ID NO 28
LENGTH: 414
TYPE: PRT
ORGANISM: saccharomyces cerevisiae
US-09-359-268A-28

Query Match      83.5%; Score 86; DB 4; Length 414;
Best Local Similarity 73.9%; Pred. No. 3.1e-08;
```

Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VEVXTVDXFGGRKDXIILSCVR 23

Db 368 IEISTVDGFGGRKDXIILSLVR 390

## RESULT 12

US-09-538-092-483  
; Sequence 483; Application US/09538092  
; Patent No. 675314  
; GENERAL INFORMATION:  
; APPLICANT: Gluc, Loic  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; CURRENT FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CuratSeqFormatter Version 0.9  
; SEQ ID NO 483  
; LENGTH: 683  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: Polypeptide Accession Number YK017C  
US-09-538-092-483

Query Match 83.5%; Score 86; DB 4; Length 683;  
Best Local Similarity 73.9%; Pred. No. 5.7e-08;

Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VEVXTVDXFGGRKDXIILSCVR 23

Db 595 IEISTVDGFGGRKDXIILSLVR 617

## RESULT 13

US-09-902-540-16148  
; Sequence 16148; Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 16148  
; LENGTH: 211  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-16148

Query Match 79.6%; Score 82; DB 4; Length 211;  
Best Local Similarity 73.9%; Pred. No. 8.3e-08;

Matches 17; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 VEVXTVDXFGGRKDXIILSCVR 23

Db 122 VEVDTVDXFGGRKDXIILSLVR 144

## RESULT 14

US-09-248-796A-19347  
; Sequence 19347; Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Kelch Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN.  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 19347  
; LENGTH: 917  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-19347

Query Match 77.7%; Score 80; DB 4; Length 917;  
Best Local Similarity 56.5%; Pred. No. 1.2e-06;  
Matches 13; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 VEVXTVDXFGGRKDXIILSCVR 23

Db 831 IMIASIDAFGRKDXIILMISCV 853

## RESULT 15

US-09-248-796A-15170  
; Sequence 15170; Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Kelch Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN.  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 15170  
; LENGTH: 426  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-15170

Query Match 76.7%; Score 79; DB 4; Length 426;  
Best Local Similarity 65.2%; Pred. No. 7.2e-07;  
Matches 15; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 VEVXTVDXFGGRKDXIILSCVR 23

Db 329 IEISTVDGFGGRKDXIILSLVR 351

Search completed: April 18, 2005, 08:18:36  
Job time : 28.708 secs

This page Blank (usps)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 18, 2005, 08:06:16 ; Search time 81.5912 Seconds  
(without alignments)  
93.693 Million cell updates/sec

Title: US-10-652-334-8  
Perfect score: 103  
Sequence: 1 VEVXTVDXFGGRKDXIILSCVR 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09C\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	94.2	23	US-10-652-334-8	Sequence 8, Appl1
2	95	92.2	380	US-10-652-334-29	Sequence 287, Ap
3	95	92.2	925	US-10-369-493-2087	Sequence 42857, A
4	94	91.3	437	US-10-425-114-22857	Sequence 57765, A
5	94	91.3	449	US-10-425-114-57765	Sequence 181690,
6	94	91.3	638	US-10-437-963-181690	Sequence 1706, Ap
7	94	91.3	734	US-10-424-599-205643	Sequence 12790, A
8	94	91.3	935	US-10-408-765A-1706	Sequence 6, Appl1
9	94	91.3	992	US-10-369-493-12790	Sequence 28, Appl
10	89	86.4	1118	US-10-474-553-6	Sequence 2003, Ap
11	86	83.5	414	US-10-652-334-28	Sequence 9, Appl1
12	86	83.5	562	US-10-104-047-2003	
13	86	83.5	677	US-10-476-924-9	

14	86	83.5	683	15	US-10-369-493-22264	Sequence 22264, A
15	83	80.6	828	15	US-10-144-194A-96	Sequence 96, Appl
16	83	80.6	2677	15	US-10-144-194A-22	Sequence 22, Appl
17	82	79.6	637	15	US-10-424-599-233501	Sequence 233501,
18	82	79.6	639	15	US-10-425-114-37717	Sequence 37717, A
19	82	79.6	1944	15	US-10-369-493-2521	Sequence 2521, Ap
20	81	78.6	642	15	US-10-369-493-21526	Sequence 21526, A
21	80	77.7	161	15	US-10-425-114-38303	Sequence 38303, A
22	80	77.7	215	15	US-10-425-114-42518	Sequence 42518, A
23	80	77.7	439	16	US-10-767-701-41078	Sequence 41078, A
24	80	77.7	653	15	US-10-369-493-21645	Sequence 21645, A
25	79	76.7	127	16	US-10-767-701-48804	Sequence 48804, A
26	79	76.7	332	15	US-10-425-114-38637	Sequence 38637, A
27	79	76.7	350	15	US-10-424-599-241211	Sequence 241211,
28	79	76.7	404	15	US-10-425-114-37833	Sequence 37833, A
29	79	76.7	626	16	US-10-437-963-166322	Sequence 166322,
30	79	76.7	642	15	US-10-425-114-37557	Sequence 37557, A
31	79	76.7	648	15	US-10-369-493-20334	Sequence 20334, A
32	79	76.7	655	16	US-10-828-924-80	Sequence 80, Appl
33	79	76.7	656	15	US-10-369-493-1268	Sequence 1268, Ap
34	79	76.7	1027	16	US-10-437-963-185291	Sequence 185291,
35	78	75.7	1223	15	US-10-424-599-216889	Sequence 216889,
36	78	75.7	611	15	US-10-369-493-10288	Sequence 10288, A
37	77	74.8	648	15	US-10-369-493-21357	Sequence 21357, A
38	76	73.8	415	16	US-10-652-334-25	Sequence 25, Appl
39	76	73.8	650	15	US-10-369-493-183044	Sequence 183044,
40	75	72.8	276	15	US-10-424-599-183044	Sequence 26, Appl
41	75	72.8	472	16	US-10-652-334-26	Sequence 1830, Ap
42	75	72.8	2231	15	US-10-369-493-1830	Sequence 21435, A
43	74	71.8	663	15	US-10-369-493-22554	Sequence 22554, A
44	74	71.8	1444	15	US-10-369-493-22554	Sequence 111794,
45	73	70.9	1323	16	US-10-437-963-111794	

#### ALIGNMENTS

RESULT 1  
US-10-652-334-8  
Sequence 8, Application US/10652334  
Publication No. US20040115787A1  
GENERAL INFORMATION:  
APPLICANT: Peltz, Stuart  
APPLICANT: Czaplinski, Kevin  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USBS  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/10/652,334  
CURRENT FILING DATE: 2003-08-28  
PRIOR APPLICATION NUMBER: US/09/359,268A  
PRIOR FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
PRIOR FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 8  
LENGTH: 23  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
FEATURE:  
OTHER INFORMATION: Xaa = any amino acid  
US-10-652-334-8

Query Match 94.2%; Score 97; DB 16; Length 23;  
Best Local Similarity 100.0%; Pred. No. 3.6e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 VEVXTVDXFGGRKDXIILSCVR 23  
DB 1 VEVXTVDXFGGRKDXIILSCVR 23

RESULT 2  
US-10-652-334-29  
; Sequence 29, Application US/10652334  
; Publication No. US20040115787A1  
; GENERAL INFORMATION:  
; APPLICANT: Pelcz, Stuart  
; APPLICANT: Czaplinski, Kevin  
; APPLICANT: Dimman, Jonathan D.  
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
; FILE REFERENCE: 601-1-85N  
; CURRENT APPLICATION NUMBER: US/10/652,334  
; CURRENT FILING DATE: 2003-08-28  
; PRIOR APPLICATION NUMBER: US/09/359,268A  
; PRIOR FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 60/093,685  
; PRIOR FILING DATE: 1998-07-22  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 29  
; LENGTH: 380  
; TYPE: PRT  
; ORGANISM: saccharomyces cerevisiae  
US-10-652-334-29

Query Match 92.2%; Score 95; DB 16; Length 380;  
Best Local Similarity 82.6%; Pred. No. 2e-08;  
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 VEVXTVDXFOGREGDXIIILSCVR 23  
|||:|||||  
Db 334 VEVASVDVAFQGREKDXIIILSCVR 356

RESULT 3  
US-10-369-493-2087  
; Sequence 2087, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2087  
; LENGTH: 925  
; TYPE: PRT  
; ORGANISM: Schizosaccharomyces pombe  
US-10-369-493-2087

Query Match 92.2%; Score 95; DB 15; Length 925;  
Best Local Similarity 82.6%; Pred. No. 5.5e-08;  
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 VEVXTVDXFOGREGDXIIILSCVR 23  
|||:|||||  
Db 749 VEVASVDVAFQGREKDXIIILSCVR 771

RESULT 4  
US-10-425-114-42857  
; Sequence 42857, Application US/10425114  
; Publication No. US20040034888A1

; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack B.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 42857  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700476463\_FLI.pep  
US-10-425-114-42857

Query Match 91.3%; Score 94; DB 15; Length 437;  
Best Local Similarity 78.3%; Pred. No. 3.5e-08;  
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 VEVXTVDXFOGREGDXIIILSCVR 23  
|||:|||||  
Db 22 IEVASVDSFOGREGDXIIILSCVR 44

RESULT 5  
US-10-425-114-57765  
; Sequence 57765, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack B.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 57765  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-GMFLMINSOY058E07\_FLI.pep  
US-10-425-114-57765

Query Match 91.3%; Score 94; DB 15; Length 449;  
Best Local Similarity 78.3%; Pred. No. 3.7e-08;  
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 VEVXTVDXFOGREGDXIIILSCVR 23  
|||:|||||  
Db 19 IEVASVDSFOGREGDXIIILSCVR 41

RESULT 6  
US-10-437-963-181690  
; Sequence 181690, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 181690  
LENGTH: 638  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(638)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_7893C.1.pep  
US-10-437-963-181690

Query Match 91.3%; Score 94; DB 16; Length 638;  
Best Local Similarity 78.3%; Pred. No. 5.4e-08;  
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEYXTVDXFGGRKDXIILSCVR 23  
DB 218 IEVASVDSFGGRKDXIILSCVR 240

RESULT 7  
US-10-424-599-205643  
Sequence 205643, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kowalc David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 205643  
LENGTH: 734  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(734)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_27723C.1.pep  
US-10-424-599-205643

Query Match 91.3%; Score 94; DB 15; Length 734;  
Best Local Similarity 78.3%; Pred. No. 6.4e-08;  
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEYXTVDXFGGRKDXIILSCVR 23  
DB 301 IEVASVDSFGGRKDXIILSCVR 323

RESULT 8  
US-10-408-765A-1706  
Sequence 1706, Application US/10408765A  
Publication No. US20040101874A1  
GENERAL INFORMATION:  
APPLICANT: Ghosh, Soumitra S.

APPLICANT: Fahy, Eoin D.  
APPLICANT: Zhang, Bing  
APPLICANT: Gibson, Bradford W.  
APPLICANT: Taylor, Steven W.  
APPLICANT: Glenn, Gary M.  
APPLICANT: Warnock, Dale E.  
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
FILE REFERENCE: 660088.465  
CURRENT APPLICATION NUMBER: US/10/408,765A  
CURRENT FILING DATE: 2003-04-04  
NUMBER OF SEQ ID NOS: 3077  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1706  
LENGTH: 935  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-408-765A-1706

Query Match 91.3%; Score 94; DB 16; Length 935;  
Best Local Similarity 78.3%; Pred. No. 8.4e-08;  
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEYXTVDXFGGRKDXIILSCVR 23  
DB 821 IEVASVDAFGGRKDXIILSCVR 843

RESULT 9  
US-10-369-493-12790  
Sequence 12790, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 12790  
LENGTH: 992  
TYPE: PRT  
ORGANISM: Aspergillus nidulans  
US-10-369-493-12790

Query Match 91.3%; Score 94; DB 15; Length 992;  
Best Local Similarity 78.3%; Pred. No. 9e-08;  
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEYXTVDXFGGRKDXIILSCVR 23  
DB 689 IEVASVDAFGGRKDXIILSCVR 711

RESULT 10  
US-10-474-553-6  
Sequence 6, Application US/10474553  
Publication No. US20040161765A1  
GENERAL INFORMATION:  
APPLICANT: JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING DISEASE  
GENES USING NONSENSE-MEDIATED DECAY INHIBITION  
FILE REFERENCE: UHV-020.25  
CURRENT APPLICATION NUMBER: US/10/474,553  
CURRENT FILING DATE: 2003-10-10  
PRIOR APPLICATION NUMBER: 60/283,920

PRIOR FILING DATE: 2001-04-13  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 1118  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-474-553-6

Query Match 86.4%; Score 89; DB 16; Length 1118;  
Best Local Similarity 77.3%; Pred. No. 8.1e-07;  
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEVXTVDXFOGREKDXIIILSCV 22  
DB 821 VEVXTVDXFOGREKDXIIILSCV 842

RESULT 11  
US-10-652-334-28

Sequence 28, Application US/10652334  
Publication No. US20040115787A1  
GENERAL INFORMATION:  
APPLICANT: Peltez, Stuart  
APPLICANT: Czaplinski, Kevin  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
FILE REFERENCE: 601-1-85N  
CURRENT FILING DATE: 2003-08-28  
PRIOR FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: US/09/359,268A  
PRIOR FILING DATE: 1998-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 28  
LENGTH: 414  
TYPE: PRT  
ORGANISM: saccharomyces cerevisiae  
US-10-652-334-28

Query Match 83.5%; Score 86; DB 16; Length 414;  
Best Local Similarity 73.9%; Pred. No. 9e-07;  
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VEVXTVDXFOGREKDXIIILSCV 23  
DB 368 IEISTVDGFOGREKDXIIILSLVR 390

RESULT 12

US-10-104-047-2003  
Sequence 2003, Application US/10104047  
Publication No. US20030236392A1  
GENERAL INFORMATION:  
APPLICANT: HELIX RESEARCH INSTITUTE  
TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA  
FILE REFERENCE: H1-A0105  
CURRENT FILING DATE: 2002-03-25  
PRIOR FILING DATE: 2002-03-25  
PRIOR APPLICATION NUMBER:  
NUMBER OF SEQ ID NOS: 4096  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2003  
LENGTH: 562  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-104-047-2003

Query Match 83.5%; Score 86; DB 15; Length 562;  
Best Local Similarity 73.9%; Pred. No. 1.3e-06;  
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VEVXTVDXFOGREKDXIIILSCV 23  
DB 446 VOVSTVDXFOGREKDXIIILSCV 468

RESULT 13  
US-10-476-924-9

Sequence 9, Application US/10476924  
Publication No. US20040152093A1  
GENERAL INFORMATION:  
APPLICANT: YUE, Henry; DING, Li;  
APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;  
APPLICANT: YUE, Huibin; HAFALIA, April J.A.;  
APPLICANT: LEE, Ernestine A.; ISON, Craig H.;  
APPLICANT: BECHA, Shanya D.; GURURAJAN, Rajagopal;  
APPLICANT: EMERLING, Brooke M.; GRIFFIN, Jennifer A.;  
APPLICANT: TANG, Y. Tom; LU, Dying Alina M.;  
APPLICANT: YAO, Montique G.; CHAMLA, Narinder K.;  
APPLICANT: RANKMAR, Jayalaxmi; GANDHI, Ameena R.;  
APPLICANT: LEE, Soo Yeun; RICHARDSON, Thomas W.;  
APPLICANT: YANG, Junming; ELIOTT, Vicki S.;  
APPLICANT: LU, Yan; THANGAVELU, Kavitha;  
APPLICANT: HE, Ann; AZIMZAI, Yalda;  
APPLICANT: RAUMANN, Brigitte E.; SWARNAKAR, Anita;  
APPLICANT: BURNFORD, Neil;  
TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS  
FILE REFERENCE: PE-0960 USN  
CURRENT FILING DATE: 2003-11-04  
PRIOR FILING DATE: 2003-11-04  
PRIOR APPLICATION NUMBER: PCT/US02/14276  
PRIOR FILING DATE: 2002-05-02  
PRIOR APPLICATION NUMBER: US 60/288,598  
PRIOR FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: US 60/291,776  
PRIOR FILING DATE: 2001-05-17  
PRIOR APPLICATION NUMBER: US 60/292,172  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/293,564  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PERL Program  
SEQ ID NO 9  
LENGTH: 677  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No: 2948827CD1  
US-10-476-924-9

Query Match 83.5%; Score 86; DB 16; Length 677;  
Best Local Similarity 73.9%; Pred. No. 1.6e-06;  
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VEVXTVDXFOGREKDXIIILSCV 23  
DB 561 VOVSTVDXFOGREKDXIIILSCV 583

RESULT 14

US-10-369-493-22264  
Sequence 22264, Application US/10369493  
Publication No. US2003023675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng



```

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO: 22264
; LENGTH: 683
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-10-369-493-22264

```

```

Query Match      83.5%; Score 86; DB 15; Length 683;
Best Local Similarity 73.9%; Pred. No. 1.6e-06;
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy      1  VEKTVDFQGRKDXIIISCV 23
      :|:|||||:|||||
Db      595 IEISTVDFQGRKDXIIISLV 617

```

```

RESULT 15
US-10-144-194A-96
; Sequence 96, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 96
; LENGTH: 828
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-144-194A-96

```

```

Query Match      80.6%; Score 83; DB 15; Length 828;
Best Local Similarity 68.2%; Pred. No. 6.9e-06;
Matches 15; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy      2  EVKTVDFQGRKDXIIISCV 23
      |||||:|||||:|||||
Db      521 EVDTVDFQGRKDXIIISCV 542

```

```

Search completed: April 18, 2005, 09:04:07
Job time : 81.5912 secs

```

This page Blank (optional)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 07:32:05 / Search time 22.1606 Seconds  
(without alignments)  
99.861 Million cell updates/sec

Title: US-10-652-334-8

Perfect score: 103

Sequence: 1 VEVXTVDXFOGQREKDXIIILSCVR 23

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: PIR 79:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	92.2	935	2	S62476
2	95	92.2	971	2	S23408
3	88	85.4	1069	2	T43280
4	86	83.5	239	2	T46441
5	86	83.5	683	2	S34700
6	82	79.6	1944	2	T40065
7	81	78.6	642	2	D69085
8	80	77.7	530	2	D70476
9	80	77.7	653	2	B75105
10	80	77.7	1090	2	T00533
11	80	77.7	1311	2	T08986
12	79	76.7	635	2	T02639
13	79	76.7	656	2	E71080
14	77	74.8	648	2	C69423
15	77	74.8	692	2	B90113
16	77	74.8	1687	2	T39072
17	76	73.8	139	2	D90303
18	76	73.8	245	2	D72258
19	76	73.8	650	2	G72429
20	76	73.8	1121	2	S50862
21	75	72.8	2231	2	S53416
22	74	71.8	663	2	H64312
23	74	71.8	993	2	A47500
24	74	71.8	1397	2	T51292
25	74	71.8	1398	2	T39568
26	72	69.9	814	2	T00740
27	72	69.9	1825	2	T52521
28	71	68.9	989	2	T48845
29	69	67.0	993	2	S35633

30	68	66.0	360	2	G64575	conserved hypothet
31	68	66.0	1077	2	T50697	hypothetical prote
32	68	66.0	2142	2	D86303	P17P16.1 protein -
33	67	65.0	1015	2	T41111	hypothetical ATP b
34	66	64.1	660	2	T41580	probable dna-bindi
35	66	64.1	1421	2	T00333	hypothetical prote
36	66	64.1	1444	2	B84809	hypothetical prote
37	65	63.1	1522	2	S48904	probable purine nu
38	65	63.1	453	2	G69494	DNA helicase homol
39	65	63.1	1075	2	C96682	protein P1B27.16
40	64	62.1	1076	2	B96682	protein P1B27.14
41	64	62.1	2219	2	T27684	hypothetical prote
42	64	61.2	1048	2	C86189	protein T25N20.11
43	62	60.2	660	2	F85069	hypothetical prote
44	62	60.2	1188	2	C71231	hypothetical prote
45	61	59.2	693	2	T26415	hypothetical prote

## ALIGNMENTS

RESULT 1  
S62476  
hypothetical protein SPAC16C9.06c - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 16-May-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S62476; T37779  
R:Badcock, K.; Churcher, C.M.  
submitted to the EMBL Data Library, October 1995  
A:Reference number: S62445  
A:Accession: S62476  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-935 <BAD>  
A:Cross-references: UNIPROT:Q09820; EMBL:Z54366; NID:q1019812; PIDN:CAA91194.1; PID:q132  
R:Badcock, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, October 1995  
A:Reference number: Z21745  
A:Accession: T37779  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 11-935 <BA2>  
A:Cross-references: EMBL:Z54366; PIDN:CAA91194.2; GSPDB:GN00066; SPDB:SPAC16C9.06c  
A:Experimental source: strain 972h; cosmid c16C9  
C:Genetics:  
A:Gene: SPDB:SPAC16C9.06c  
A:Map position: 1L

Query Match 92.2%; Score 95; DB 2; Length 935;  
Best Local Similarity 82.6%; Pred. No. 1.2e-08;  
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEVXTVDXFOGQREKDXIIILSCVR 23  
DB 759 VEVASVDAFOGQREKDXIIILSCVR 781

RESULT 2  
S23408  
prematurely terminated mRNA decay factor NAM7 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein YM9583.05c; protein YMR080c; UPF1 protein  
C:Species: Saccharomyces cerevisiae  
C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: S23408; A44386; S54455  
R:Altamura, N.; Groudinsky, O.; Dujardin, G.; Stonitski, P.P.  
J. Mol. Biol. 224, 575-587, 1992  
A:Title: NAM7 nuclear gene encodes a novel member of a family of helicases with a Zn-11g  
A:Reference number: S23408; MUID:22235815; PMID:11314899  
A:Accession: S23408  
A:Molecule type: DNA  
A:Residues: 1-971 <ALT>  
A:Cross-references: UNIPROT:P30771; EMBL:X62394; NID:q4022; PIDN:CAA44266.1; PID:q4023  
R:leeds, P.; Wood, J.M.; Lee, B.S.; Gilbertson, M.R.

Mol. Cell. Biol. 12, 2165-2177, 1992.  
A:Title: Gene products that promote mRNA turnover in *Saccharomyces cerevisiae*.  
A:Reference number: A44388; MUID:92236591; PMID:1569946  
A:Accession: A44388  
A:Molecule type: DNA  
A:Residues: 1-971 <LEB>  
A:Cross-references: GB:M76659; NID:G173141; PIDN:AAA5197.1; PID:G173142  
A:Gentiles, S.; Bowman, S.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: S54451  
A:Accession: S54455  
A:Molecule type: DNA  
A:Residues: 1-971 <GEN>  
A:Cross-references: EMBL:Z49259; NID:G807956; PID:G807962; MIPS:YMR080C  
A:Experimental source: strain AB972  
C:Genetics:  
A:Gene: SGD:NAM7; UPR1  
A:Cross-references: SGD:S0004685; MIPS:YMR080C  
A:Map position: 13R  
C:Keywords: GTP binding; mitochondrial; nucleotide binding; nucleus; P-loop  
P:430-437/Region: nucleotide-binding motif A (P-loop)  
P:545-548/Region: GTP-binding NKXD motif

Query Match 92.2%; Score 95; DB 2; Length 971;  
Best Local Similarity 82.6%; Pred. No. 1.2e-08;  
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 VEVXTVDXFOGREKDXIIISLSCVR 23  
DB 757 VEVASVDAFOGREKDXIIISLSCVR 779

## RESULT 3

T43280  
nonsense-mediated mRNA decay trans-acting factor - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T43280  
R:Page, M.F.; Carr, B.; Anders, K.R.; Gilmson, A.; Anderson, P.  
Mol. Cell. Biol. 19, 5943-5951, 1999  
A:Title: SMG-2 is a phosphorylated protein required for mRNA surveillance in *Caenorhabditis elegans*  
A:Reference number: Z22389; MUID:99384262; PMID:10454541  
A:Accession: T43280  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1069 <PAG>  
A:Cross-references: UNIPROT:O76512; EMBL:AF074017; NID:G3328176; PIDN:AAC26789.1; PID:G3

Query Match 85.4%; Score 88; DB 2; Length 1069;  
Best Local Similarity 69.6%; Pred. No. 2.4e-07;  
Matches 16; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 VEVXTVDXFOGREKDXIIISLSCVR 23  
DB 795 VEVASVDAFOGREKDXIIISLSCVR 817

## RESULT 4

T46441  
hypothetical protein DKFZp434C0927.1 - human  
C:Species: *Homo sapiens* (man)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004  
C:Accession: T46441  
R:Bioecker, H.; Beecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z23032  
A:Accession: T46441  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-239 <AAA>  
A:Cross-references: UNIPROT:Q9NSW3; EMBL:AL137700  
A:Experimental source: adult testis; clone DKFZp434C0927  
C:Genetics:

A>Note: DKFZp434C0927.1

Query Match 83.5%; Score 86; DB 2; Length 239;  
Best Local Similarity 73.9%; Pred. No. 1.3e-07;  
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 VEVXTVDXFOGREKDXIIISLSCVR 23  
DB 123 VQVSTVDXFOGREKDXIIISLSCVR 145

## RESULT 5

S34700  
probable purine nucleotide-binding protein YKL017C - yeast (*Saccharomyces cerevisiae*)  
C:Species: *Saccharomyces cerevisiae*  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: S34700; S37830; S37834  
R:Wiemann, S.; Voss, H.; Schwaeger, C.; Rupp, T.; Stegemann, J.; Zimmermann, J.; Grothues, A.; Descriptions: Sequencing and analysis of 51.5 kilobases on the left arm of chromosome XI  
A:Reference number: S34679  
A:Accession: S34700  
A:Molecule type: DNA  
A:Residues: 1-683 <WLB>  
A:Cross-references: UNIPROT:P34243; EMBL:X74152; NID:G450363; PID:G395256  
A:Experimental source: strain S288C  
R:Wiemann, S.; Voss, H.; Schwaeger, C.; Rupp, T.; Grothues, D.; Sensen, C.; Stegemann, J.  
submitted to the Protein Sequence Database, March 1994

Query Match 83.5%; Score 86; DB 2; Length 683;  
Best Local Similarity 73.9%; Pred. No. 3.5e-07;  
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 VEVXTVDXFOGREKDXIIISLSCVR 23  
DB 595 IEISTVDXFOGREKDXIIISLSCVR 617

## RESULT 6

T40065  
tRNA-splicing endonuclease positive effector - fission yeast (*Schizosaccharomyces pombe*)  
C:Species: *Schizosaccharomyces pombe*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T40065  
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Delaure, V.; Gallibert, F.  
submitted to the EMBL Data Library, December 1998  
A:Reference number: Z21903  
A:Accession: T40065  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1944 <LYN>

Query Match 83.5%; Score 86; DB 2; Length 683;  
Best Local Similarity 73.9%; Pred. No. 3.5e-07;  
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 VEVXTVDXFOGREKDXIIISLSCVR 23  
DB 595 IEISTVDXFOGREKDXIIISLSCVR 617

A:Cross-references: UNIPROT:O94387; EMBL:AL034463; PIDN:CAA22438.1; GSPDB:GN00067; SPDB:  
A:Experimental source: strain 972h-; cosmid c29A10  
C:Genetics:  
A:Gene: SPDB:SPBC29A10.10c  
A:Map position: 2

Query Match 79.6%; Score 82; DB 2; Length 1944;  
Best Local Similarity 65.2%; Pred. No. 5.1e-06;  
Matches 15; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VEVXTVDXFOGREGKXIIISLCVR 23  
Db 1700 LDHVTDFGQGRKXIIISFCVR 1722

RESULT 7  
D69085  
transcription control factor enhancer-binding protein - Methanobacterium thermoautotroph  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: D69085  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
; Qiu, D.; Spadofora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.  
K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func  
A:Reference number: A69000; MUID:98037514; PMID:9371463  
A:Accession: D69085  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-642 <MT>  
A:Cross-references: UNIPROT:O27671; GB:AE000922; GB:AE000666; NID:g2622754; PIDN:AA8610  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH1634  
C:Superfamily: probable DNA helicase MJ0104

Query Match 78.6%; Score 81; DB 2; Length 642;  
Best Local Similarity 73.9%; Pred. No. 2.6e-06;  
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VEVXTVDXFOGREGKXIIISLCVR 23  
Db 551 VEVNSVDGFOGREGKXIIISLVR 573

RESULT 8  
D70476  
DNA helicase - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C:Accession: D70476  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O  
V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: D70476  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-530 <AOP>  
A:Cross-references: UNIPROT:O67840; GB:AB000770; NID:g2984274; PIDN:AAC07803.1; PID:g298  
A:Experimental source: strain VFS  
C:Genetics:  
A:Gene: helix

Query Match 77.7%; Score 80; DB 2; Length 530;  
Best Local Similarity 73.9%; Pred. No. 3.3e-06;  
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VEVXTVDXFOGREGKXIIISLCVR 23  
Db 450 VEVKTVDGFOGREGKXIIISLVR 472

RESULT 9  
B75105  
probable DNA helicase PAB1561 - Pyrococcus abyssi (strain Orsay)  
C:Species: Pyrococcus abyssi  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: B75105  
R:anonymous; Genoscope  
submitted to the EMBL Data Library, July 1999  
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru  
A:Reference number: A75001  
A:Accession: B75105  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-653 <XAM>  
A:Cross-references: UNIPROT:Q9UBZ6; GB:AJ248286; GB:AL096836; NID:g5458366; PIDN:CAB5014  
A:Experimental source: strain Orsay  
C:Genetics:  
A:Gene: PAB1561  
C:Superfamily: probable DNA helicase MJ0104

Query Match 77.7%; Score 80; DB 2; Length 653;  
Best Local Similarity 73.9%; Pred. No. 4e-06;  
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VEVXTVDXFOGREGKXIIISLCVR 23  
Db 570 VEVKTVDGFOGREGKXIIISLVR 592

RESULT 10  
T00533  
probable DNA2-NAM7 helicase family protein [imported] - Arabidopsis thaliana  
N:Alternate names: SSN1 protein homolog T20K24.14  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: T00533; G84572  
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Kechum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,  
submitted to the EMBL Data Library, July 1997  
A:Description: Arabidopsis thaliana chromosome II BAC T20K24 genomic sequence.  
A:Reference number: Z14167  
A:Accession: T00533  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1090 <ROU>  
A:Cross-references: UNIPROT:O64476; EMBL:AC002392; NID:g3176701; PID:g3176714  
A:Experimental source: cultivar Columbia  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentol, M.L.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Crohn, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: G84572  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1090 <STO>  
A:Cross-references: GB:AE002093; NID:g3176714; PIDN:AA012029.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: T20K24.14; AC2g19120  
A:Map position: 2  
A:introns: 250/1; 310/1; 342/1; 383/2; 427/3; 1016/3

Query Match 77.7%; Score 80; DB 2; Length 1090;  
Best Local Similarity 60.9%; Pred. No. 6.6e-06;  
Matches 14; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VEVXTVDXFOGREGKXIIISLCVR 23  
Db 967 IYINTVDGFOGREGKXIIISLVR 989

RESULT 11  
T08986  
hypotheoretical protein F6G3.130 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C/Accession: T08986  
R/Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X  
submitted to the Protein Sequence Database, May 1999  
A/Reference number: Z16520  
A/Accession: T08986  
A/Molecule type: DNA  
A/Residues: 1-1311 <BEV>  
A/Cross-references: UNIPROT:Q9SZW3; EMBL:AL078464; GSPDB:GN00062; ATSP:F6G3.130  
C/Experimental source: cultivar Columbia; BAC clone F6G3  
C/Genetics:  
A/Gene: ATSP:F6G3.130  
A/Map position: 4  
A/Intons: 414/1; 474/1; 506/1; 547/2; 591/3; 1179/3  
Query Match 77.7%; Score 80; DB 2; Length 1311;  
Best Local Similarity 60.9%; Pred. No. 7.9e-06;  
Matches 14; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 VEVXTVDXFGQREKDXIIISGCR 23  
: ||| |||: |||: |||  
Db 1130 IYINTVDAFGQERDVIIISGCR 1152

RESULT 12  
T02659  
Probable helicase At2g03270 (imported) - Arabidopsis thaliana  
N/Alternate names: hypotheoretical protein T18E12.6  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C/Accession: T02659; D84446  
R/Runnaley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Rot  
submitted to the EMBL Data Library, September 1998  
A/Description: Arabidopsis thaliana chromosome II BAC T18E12 genomic sequence.  
A/Reference number: Z14702  
A/Accession: T02659  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-635 <ROU>  
A/Cross-references: UNIPROT:O81047; EMBL:AC005313; NID:G3548797; PID:G3548803  
A/Experimental source: cultivar Columbia  
R/Lin, X.; Kaul, S.; Runnaley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Unayam, L.; Tallon, L.;  
euser, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MUID:20083487; PMID:10617197  
A/Accession: D84446  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-635 <STD>  
A/Cross-references: GB:AE002093; NID:G4335770; PID:AAU17447.1; GSPDB:GN00139  
C/Genetics:  
A/Gene: At2g03270; T18E12.6  
A/Map position: 2  
C/Superfamily: probable DNA helicase MJ0104  
Query Match 76.7%; Score 79; DB 2; Length 635;  
Best Local Similarity 65.2%; Pred. No. 5.9e-06;  
Matches 15; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 VEVXTVDXFGQREKDXIIISGCR 23  
: ||| |||: |||: |||  
Db 550 MEISTVDGFGQREKAXIIISVR 572

RESULT 13  
E71080  
probable DNA-binding protein - Pyrococcus horikoshii

C/Species: Pyrococcus horikoshii  
C/Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 12-Jul-2004  
C/Accession: E71080  
R/Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine  
M.; Ohkubo, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi,  
DNA Res. 5, 55-76, 1998  
A/Title: Complete sequence and gene organization of the genome of a hyper-thermophilic ar  
A/Reference number: A71000; MUID:38344137; PMID:9679194  
A/Accession: E71080  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-656 <KAW>  
A/Cross-references: UNIPROT:O58624; GB:AP000004; NID:G3236131; PID:BA330003.1; PID:G325;  
A/Experimental source: strain OT3  
A/Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C/Genetics:  
A/Gene: PH0909

Query Match 76.7%; Score 79; DB 2; Length 656;  
Best Local Similarity 73.9%; Pred. No. 6.1e-06;  
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VEVXTVDXFGQREKDXIIISGCR 23  
: ||| |||: |||: |||  
Db 570 VEVXTVDGFGQREKXVIIISFVR 592

RESULT 14  
C69423  
DNA helicase homolog - Archaeoglobus fulgidus  
C/Species: Archaeoglobus fulgidus  
C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C/Accession: C69423  
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.R.; Ketchum, K.A.; Dodson,  
P.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirsch, B.F.;  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A/Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.J.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeor  
A/Reference number: A69250; MUID:98049343; PMID:9389475  
A/Accession: C69423  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-648 <KLE>  
A/Cross-references: UNIPROT:Q28883; GB:AE001009; GB:AE000782; NID:G2689332; PID:AA88986C  
C/Superfamily: probable DNA helicase MJ0104

Query Match 74.8%; Score 77; DB 2; Length 648;  
Best Local Similarity 69.6%; Pred. No. 1.4e-05;  
Matches 16; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 VEVXTVDXFGQREKDXIIISGCR 23  
: ||| |||: |||: |||  
Db 556 VEVSSVDGFGQREKXVIIISFVR 578

RESULT 15  
E90113  
hypotheoretical protein component of a tRNA splicing complex (imported) - Guillardia theta r  
C/Species: nucleomorph Guillardia theta  
A/Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C/Accession: E90113  
R/Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reil  
Nature 410, 1091-1096, 2001  
A/Title: The highly reduced genome of an enslaved algal nucleus.  
A/Reference number: A99082; MUID:11323671; PMID:11323671  
A/Accession: E90113  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-692 <DOU>  
A/Cross-references: UNIPROT:Q9AVZ7; GB:AO10592; NID:G12580756; PID:CAC27074.1; GSPDB:GN

C;Genetics:  
 A;Gene: component of a tRNA splicing complex  
 A;Map position: 2  
 A;Genome: nucleomorph  
 C;Keywords: nucleomorph

Query Match 74.8%; Score 77; DB 2; Length 692;  
 Best Local Similarity 54.5%; Pred. No. 1.5e-05;  
 Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 EVTVDXFGREKDXIIISCVR 23  
 Db 570 QISTIDFQGRKDIILFSCVR 591

Search completed: April 18, 2005, 08:06:05  
 Job time : 23.1606 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 07:33:55 ; Search time 96.0292 Seconds

(without alignments)  
122.648 Million cell updates/sec

Title: US-10-652-334-8

Sequence: 1 VEVXVTVDXFGQREKMDXIIILSCVR 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Uniprot\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95	92.2	925	1 RNT1 SCHPO	O09820 schizosacch
2	95	92.2	964	2 O6FVZ1	O6FVZ1 candida gla
3	95	92.2	969	1 O6CWC8	O6CWC8 kluyveromyc
4	95	92.2	971	1 NAM7 YEAST	P30771 saccharomyc
5	95	92.2	1000	2 O75DS7	O75DS7 aabhya gos
6	94	91.3	543	2 O8K0N4	O8K0N4 mus musculu
7	94	91.3	985	2 O6BPM3	O6BPM3 debaryomyc
8	94	91.3	1060	2 O8TFW3	O8TFW3 aspergillus
9	94	91.3	1079	2 O6MYT2	O6MYT2 aspergillus
10	94	91.3	1093	1 RNT1 NEUCR	O9NEH1 neurospora
11	94	91.3	1097	1 RNT1 FUGRU	O9ETX3 fugu rubrip
12	94	91.3	1098	2 O6GNR2	O6GNR2 xenopus lae
13	94	91.3	1100	2 O7ZVZ4	O7ZVZ4 brachydanio
14	94	91.3	1113	1 RNT1 MOUSE	O9EPU8 mus musculu
15	94	91.3	1113	2 O6GYF5	O6GYF5 mus musculu
16	94	91.3	1118	2 O6BZ25	O6BZ25 homo sapien
17	94	91.3	1124	2 O6PHQ5	O6PHQ5 mus musculu
18	94	91.3	1129	1 RNT1 HUMAN	O9J900 homo sapien
19	94	91.3	1235	1 RNT1 ARATH	O9JF90 arabidopsis
20	94	91.3	1243	2 O8S3K7	O8S3K7 arabidopsis
21	94	91.3	1277	2 O6SVT5	O6SVT5 oryza sativ
22	92	89.3	1297	2 O7RQ16	O7RQ16 plasmodium
23	92	89.3	1554	2 O8IJY4	O8IJY4 plasmodium
24	92	89.3	1554	2 O6C803	O6C803 yarrowia li
25	91	88.3	1120	2 O7PMW4	O7PMW4 anopheles g
26	91	88.3	1180	1 RNT1 DROME	O9YV83 drosophila
27	88	85.4	1069	1 RNT1 CAEBL	O7F512 caenorhadi
28	86	83.5	239	2 O9NSW3	O9NSW3 homo sapien
29	86	83.5	527	2 O6CHW3	O6CHW3 yarrowia li
30	86	83.5	683	1 YKB7 YEAST	P34243 saccharomyc
31	86	83.5	926	2 O6ZU11	O6ZU11 homo sapien

32	86	83.5	1139	2 O7RRP6	O7RRP6 plasmodium
33	85	82.5	649	2 O6CFH6	O6CFH6 yarrowia li
34	84	81.6	770	2 O7X684	O7X684 oryza sativ
35	84	81.6	778	2 O6Z081	O6Z081 mus musculu
36	84	81.6	821	2 O6PED8	O6PED8 mus musculu
37	84	81.6	902	2 O8OV90	O8OV90 mus musculu
38	84	81.6	2646	2 O6IMG6	O6IMG6 mus musculu
39	84	81.6	2743	2 O8IER9	O8IER9 plasmodium
40	83	80.6	917	2 O6AZD7	O6AZD7 homo sapien
41	83	80.6	2677	1 SETX HUMAN	O7Z333 homo sapien
42	82	79.6	755	2 O6BKZ7	O6BKZ7 debaryomyc
43	82	79.6	1944	2 O94387	O94387 schizosacch
44	81	78.6	237	2 O8K2R9	O8K2R9 mus musculu
45	81	78.6	528	2 O8BQV5	O8BQV5 m mus musculu

## ALIGNMENTS

RESULT 1  
RNT1 SCHPO STANDARD; PRT; 925 AA.  
ID RNT1 SCHPO  
AC O09820;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Regulator of nonsense transcripts 1 homolog.  
GN ORFNames=SPAC16C9.06c;  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;  
RX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Spouris J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,  
RA Gentles S., Gobie A., Hamlin N., Harris N., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odel C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Voicakeert G., Aert R., Robben J., Grympey B.,  
RA Welljens I., Vansteede E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moser D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Leinach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Parnelle B.,  
RA Goffeau A., Cadieu B., Diano S., Gloux S., Lelaure V., Mottler S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas R., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Moreno S., Armstrong J., Forsburg S.L.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Uesery D., Barrell B.G., Nure P.,  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
- FUNCTION: Eliminates the production of nonsense-containing RNAs.  
- SUBCELLULAR LOCATION: Cytoplasmic (potential).  
- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.  
- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@eb-sib.ch).

CC EMBL; Z54366; CA91194.2; -.  
 DR PIR; S62476; S62476.  
 DR GenBank; SP063593; AAA ATase.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR006935; Resili.  
 DR SMART; SM00382; AAA; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR ATP-binding; Helicase; Hypothetical protein;  
 KM Nonsense-mediated mRNA decay; Zinc-finger.  
 FT ZN FING 52 80 C2H2-type (atypical) (Potential).  
 FT ZN FING 104 134 C4-type (Potential).  
 FT NP BIND 414 421 ATP (By similarity).  
 SQ SEQUENCE 925 AA; 104528 MW; 4A5D63C828E864 CRC64;

Query Match 92.2%; Score 95; DB 1; Length 925;  
 Best Local Similarity 82.6%; Pred. No. 1.3e-08;  
 Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VEVXTVDXFGREKDXIILSCVR 23

Db 749 VEVASVDAFQGREKDXIILSCVR 771

RESULT 2

Q6FVZ1 PRELIMINARY; PRT; 964 AA.

AC Q6FVZ1; 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DN Candida glabrata strain CBS138 chromosome D complete sequence.  
 OS ORFNames=CAGJ0D043129;  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; mitospotic Saccharomycetales; Candida.  
 NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RA NCBI\_TaxID=284593;

RESULT 3

Q6CW68 PRELIMINARY; PRT; 969 AA.

AC Q6CW68; 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DN Kluyveromyces lactis strain NRRL Y-1140 chromosome B of strain NRRL Y-1140 of Kluyveromyces lactis.  
 OS ORFNames=KLA0B064359;  
 CC Kluyveromyces lactis NRRL Y-1140.  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.  
 NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

RA NCBI\_TaxID=284590;

Query Match 92.2%; Score 95; DB 2; Length 969;  
 Best Local Similarity 82.6%; Pred. No. 1.4e-08;  
 Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VEVXTVDXFGREKDXIILSCVR 23

Db 755 VEVASVDAFQGREKDXIILSCVR 777

RESULT 4

NAM7 YEAST STANDARD; PRT; 971 AA.

AC P30771; 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-JUL-1993 (Rel. 26, Last annotation update)  
 DN NAM7 protein (Nuclear accommodation of mitochondria 7 protein)  
 DE Nonsense-mediated mRNA decay protein 1 (Up-frameshift suppressor 1).  
 GN Name=NAM7; Synonyms=IFS2, MOR4, UPFI; OrderedNames=YM8080C;  
 OS Saccharomycetes cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 NCBI\_TaxID=4932;

RA NCBI\_TaxID=4932;

RA NCBI\_TaxID=4932;

RA NCBI\_TaxID=4932;

RA NCBI\_TaxID=4932;

RA NCBI\_TaxID=4932;

RA NCBI\_TaxID=4932;

RA NCBI\_TaxID=4932;

RA NCBI\_TaxID=4932;

RA NCBI\_TaxID=4932;

RP SEQUENCE FROM N.A.  
 RC STRAIN=R23/50;  
 RA MEDLINE=92235815; PubMed=1314899;  
 RA Altamura N., Grudinsky O., Dujardin G., Slonimski P.P.;  
 RT "NAM7 nuclear gene encodes a novel member of a family of helicases  
 RT with a Zn-ligand motif and is involved in mitochondrial functions in  
 RT Saccharomyces cerevisiae.";  
 RL J. Mol. Biol. 224:575-587(1992).  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92235591; PubMed=1569946;  
 RA Leeds P.F., Wood J.M., Lee B.S., Culbertson M.R.;  
 RT "Gene products that promote mRNA turnover in Saccharomyces  
 RT cerevisiae.";  
 RL Mol. Cell. Biol. 12:2165-2177(1992).  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RX MEDLINE=97313268; PubMed=9169872;  
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,  
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,  
 RA Ugeux K., Lye G., Moulé S., Odell C., Pearson D., Rajandream M.A.,  
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrall B.G.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome  
 RT XII.";  
 RL Nature 387:90-93(1997).  
 RN  
 RP CHARACTERIZATION.  
 RX MEDLINE=97051830; PubMed=8896465;  
 RA Cui Y., Dimman J.D., Peltz S.W.;  
 RT "Mof4-1 is an allele of the UPF1/IPS2 gene which affects both mRNA  
 RT turnover and -1 ribosomal frameshifting efficiency.";  
 RL EMBL J. 15:5726-5736(1996)  
 CC  
 CC -1- FUNCTION: Probable helicase involved in mitochondrial functions.  
 CC Required for rapid turnover of mRNAs containing a premature  
 CC translational termination codon.  
 CC -1- SUBCELLULAR LOCATION: Present predominantly in the cytoplasm, but  
 CC is also found in small quantities in the nucleus.  
 CC  
 CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR EMBL; X62394; CAA44266.1; -;  
 DR EMBL; M76659; AAA5197.1; -;  
 DR EMBL; Z49259; CAA89226.1; -;  
 DR PIR; S23408; S23408.  
 DR GenBank; 142747; -;  
 DR SGD; S000004685; NAM7.  
 DR GO; GO:0005737; C:cytoplasm; IDA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR001410; DEAD.  
 DR SMART; SM00382; AAA; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 KM ATP-binding; Helicase; Hydrolyase; Mitochondrion;  
 KM Nonsense-mediated mRNA decay; Nuclear protein; Zinc-finger.  
 FT ZN\_FING 70 98 C2H2-type (atypical) (Potential).  
 FT NP\_BIND 122 152 C4-type (Potential).  
 FT NP\_BIND 430 437 ATP (By similarity).  
 SQ SEQUENCE 971 AA; 109430 MW; 9161AFB0BE6747FE CRC64;  
 Query March 92.2%; Score 95; DB 1; Length 971;  
 Best Local Similarity 82.6%; Pred. No. 1.4e-08;  
 Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RESULT 5  
 ID 075DS7 PRELIMINARY; PRT; 1000 AA.  
 AC 075DS7;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE ABR022CP.  
 GN ORFNames=ABR022C;  
 OS Asibhya gossypii (Yeast) (Eremothecium gossypii).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Eremothecium.  
 CC NCBI\_TaxId=33169;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 10895;  
 RA Voegeli S.E., Dietrich F.S., Brachat S., Lerch A., Gaffney T.,  
 RA Philippson P.;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB016815; AAS50792.2; -;  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR SMART; SM00382; AAA; 1.  
 KM ATP-binding.  
 SQ SEQUENCE 1000 AA; 111963 MW; 39D4257F675B72A6 CRC64;  
 Query March 92.2%; Score 95; DB 2; Length 1000;  
 Best Local Similarity 82.6%; Pred. No. 1.4e-08;  
 Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RESULT 6  
 ID 08K0N4 PRELIMINARY; PRT; 543 AA.  
 AC 08K0N4;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Renti protein (Fragment).  
 GN Name=Renti;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxId=10090;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Eye;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhac N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsten F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshitsuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Vallalath D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalins D.E., Schenck A., Schein J.E.,  
 RA Jones S.J., Maier M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bye;  
 RA Strauberg R.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC030916; AAH0916.1; -.  
 DR MGD; MGI:107995; Renti1.  
 FT NON TER 1  
 SQ SEQUENCE 543 AA; 59832 MW; B395487F83CB0E57 CRC64;

Query Match 91.3%; Score 94; DB 2; Length 543;  
 Best Local Similarity 78.3%; Pred. No. 1.1e-08;  
 Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEVXTVDXFOGREKDXIIILSCVR 23  
 DB 246 VEIVSVDAPFOGREKDXIIILSCVR 268

## RESULT 7

06BPM3 PRELIMINARY; PRT; 985 AA.  
 AC 06BPM3;  
 DT 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE Debaryomyces hanseni chromosome E of strain CBS767 of Debaryomyces hanseni.  
 GN ORFNames=DEHA0E13002g;  
 OS Debaryomyces hanseni CBS767.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Debaryomyces.  
 NX NCBI\_TaxID=284592;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS767;  
 RA Genolevures;  
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S., Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla B., Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V., Barnay S., Blanchin S., Beckerich J.M., Beyne B., Bleykasten C., Boissarme A., Boyer J., Cattolico L., Confanioli F., de Darnvar A., Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A., Hantay F., Hennequin C., Jauniaux N., Joyet P., Kachouri R., Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H., Nicoud J.M., Nikolaki M., Ozias S., Ozier-Kalogeropoulos O., Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A., Sweeney D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B., Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A., RA Boucher C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J., RA Wincker P., Soulet J.L.;  
 RT "Genome evolution in yeasts."  
 RL Nature 430:35-44(2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS767;  
 RA Genoscope;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR82137; CAG8086.1; -.  
 DR InterPro; IPR011545; DEAD/DEAH\_N.  
 DR SMART; SM00487; DEXDC; 1.  
 SQ SEQUENCE 985 AA; 110705 MW; 14BBD331A4E37E0D CRC64;

Query Match 91.3%; Score 94; DB 2; Length 985;  
 Best Local Similarity 78.3%; Pred. No. 2.2e-08;  
 Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEVXTVDXFOGREKDXIIILSCVR 23  
 DB 784 VEIVSVDAPFOGREKDXIIILSCVR 806

## RESULT 8

08TFW3 PRELIMINARY; PRT; 1060 AA.  
 AC 08TFW3;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Possible regulator of nonsense transcripts.  
 GN Name=AF5C11.22c;  
 OS Aspergillus fumigatus (Sartorya fumigata).  
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
 NX NCBI\_TaxID=5085;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AF293;  
 RA Knowles D.G., Warren T., Hall N., Quail M., Woodward J.R., Denning D.W., Anderson M.J., Bartell B.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL713629; CAD28448.1; -.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:000166; F:nucleotide binding; IEA.  
 DR GO; GO:0015668; F:type III site-specific deoxyribonuclease ac. . .; IEA.  
 DR GO; GO:0009307; P:DNA restriction; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR011545; DEAD/DEAH\_N.  
 DR InterPro; IPR006935; Resili.  
 DR Pfam; PF04851; Resili; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 KW ATP-binding.  
 SQ SEQUENCE 1060 AA; 116728 MW; 744DFCS8A26E57B CRC64;

Query Match 91.3%; Score 94; DB 2; Length 1060;  
 Best Local Similarity 78.3%; Pred. No. 2.4e-08;  
 Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEVXTVDXFOGREKDXIIILSCVR 23  
 DB 757 VEIVSVDAPFOGREKDXIIILSCVR 779

## RESULT 9

06MY12 PRELIMINARY; PRT; 1079 AA.  
 AC 06MY12;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Regulator of nonsense transcripts, putative.  
 GN ORFNames=AF5C11.22c;  
 OS Aspergillus fumigatus (Sartorya fumigata).  
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
 NX NCBI\_TaxID=5085;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=14998527; DOI=10.1016/j.fgb.2003.12.003;  
 RA Pain A., Woodward J., Quail M.A., Anderson M.J., Clark R., Collins M., Foster N., Fraser A., Harris D., Latke N., Murphy L., Humphray S., O'Neill S., Petrea M., Price C., Rabdinowitsch E., Rajandream M.-A., Salzberg S., Saunders B., Seegar K., Sharp S., Warren T., RA Denning D.W., Bartell B., Hall N.;  
 RT "Insight into the genome of Aspergillus fumigatus: analysis of a 922 kb region encompassing the nitrate assimilation gene cluster."  
 RL Fungal Genet. Biol. 41:443-453(2004).  
 DR EMBL; BX649606; CAF32021.1; -.  
 DR InterPro; IPR011545; DEAD/DEAH\_N.  
 DR InterPro; IPR006935; Resili.  
 DR Pfam; PF04851; Resili; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 SQ SEQUENCE 1079 AA; 118655 MW; A8907E6FA8500E63 CRC64;

Query Match 91.3%; Score 94; DB 2; Length 1079;  
 Best Local Similarity 78.3%; Pred. No. 2.4e-08;  
 Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEVXVDXFGREKDXIIISLSCVR 23

DB 776 EVASVDAFQGRKDFIIVLSCVR 798

RESULT 10  
 RNT1 NEUCR STANDARD; PRT; 1093 AA.  
 AC Q9HEH1; Q7RVU9;  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Regulator of nonsense transcripts 1 homolog.  
 GN ORFNames=284.130, NCU04242.1;  
 OS Neurospora crassa.  
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 CC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 CC NCBI\_TaxID=5111;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=74-OR23-1A / FGSC 987;  
 RC PubMed=12712197; DOI=10.1038/nature01554;  
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
 RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehan B.,  
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Qui D., Janakiev P., Bell-Pedersen D., Nelson M.A.,  
 RA Werner-Washburne M., Selitrenikoff C.P., Kinsey J.A., Braun E.L.,  
 RA Zelter A., Schulte U., Kothe G.O., Jedd G., Mewes H.-W., Staben C.,  
 RA Marcotte B., Greenberg D., Roy A., Foley K., Naylor J.,  
 RA Stange-Thomann N., Barrett R., Gnerre S., Kamal M., Kamysovs M.,  
 RA Mancel E., Bielke C., Rudd S., Frisman D., Krystofova S.,  
 RA Rasmussen C., Weisenberg R.L., Perkins D.D., Kroken S., Cogoni C.,  
 RA Desouza C.P., Glass L., Orbach M.J., Berglund J.A., Voelker R.,  
 RA Varden O., Plamann M., Selter S., Dunlap J., Radford A., Aramayo R.,  
 RA Naveg D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nudbaum C., Birren B.,  
 RA "The genome sequence of the filamentous fungus Neurospora crassa."  
 RL Nature 423:859-868(2003).  
 CC -1- FUNCTION: Eliminates the production of nonsense-containing RNAs  
 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- SIMILARITY: Belongs to the Dna2/Nam7 helicase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL451023; CAC18314.1; -;  
 DR EMBL; AABX01000272; EAA11997.1; -;  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR006935; RecIII.  
 DR SMART; SM00382; AAA; 1.  
 DR SMART; SM00487; DEXDC; 1.

KM ATP-binding; Helicase; Hydrolase; Hypothetical protein;  
 KM Nonsense-mediated mRNA decay; Zinc-finger.  
 FT ZN FING 111 139 C2H2-type (atypical) (Potential).  
 FT ZN FING 163 193 C4-type (Potential).  
 FT NP\_BIND 477 484 ATP (Potential).  
 FT DOMAIN 59 62 Poly-Asp.  
 FT DOMAIN 69 73 Poly-Asp.  
 FT SEQUENCE 1093 AA; 120087 MW; 8B0E4F047ACB142 CRC64;

Query Match 91.3%; Score 94; DB 1; Length 1093;  
 Best Local Similarity 78.3%; Pred. No. 2.5e-08;  
 Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEVXVDXFGREKDXIIISLSCVR 23

DB 806 EVASVDAFQGRKDFIIVLSCVR 828

RESULT 11  
 RNT1 FUGRU STANDARD; PRT; 1097 AA.  
 AC Q98TR3;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Putative regulator of nonsense transcripts 1.  
 GN Name=RENT1;  
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Neuteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neuteleostei;  
 CC Acanthopterygii; Acanthopterygii; Perciformes; Tetraodontiformes;  
 CC Tetraodontidae; Tetraodontidae; Takifugu.  
 CC NCBI\_TaxID=31033;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Clarke D., Elgar G., Clark M.S.;  
 RT "Comparative analysis of human 19p12-13 region in Fugu and mouse."  
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: Eliminates the production of nonsense-containing RNAs  
 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- SIMILARITY: Belongs to the Dna2/Nam7 helicase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ301641; CAC33025.1; -;  
 DR InterPro; IPR006935; RecIII.  
 KM ATP-binding; Helicase; Hydrolase; Nonsense-mediated mRNA decay;  
 KM Zinc-finger.  
 FT ZN FING 113 141 C2H2-type (atypical) (Potential).  
 FT ZN FING 165 195 C4-type (Potential).  
 FT NP\_BIND 474 481 ATP (Potential).  
 FT SEQUENCE 1097 AA; 122338 MW; FDC7C2B4E26AC54 CRC64;

Query Match 91.3%; Score 94; DB 1; Length 1097;  
 Best Local Similarity 78.3%; Pred. No. 2.5e-08;  
 Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEVXVDXFGREKDXIIISLSCVR 23

DB 801 EVASVDAFQGRKDFIIVLSCVR 823

RESULT 12  
 O6GNR2 PRELIMINARY; PRT; 1098 AA.  
 AC O6GNR2;  
 AC O6GNR2;

```

DT 05-JUN-2004 (TREMBlrel. 27, Created)
DT 05-JUN-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUN-2004 (TREMBlrel. 27, Last annotation update)
DE MCG80941 protein.
GN Name=MCG80941; (African clawed frog).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT Initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC073441; AAH73441.1;
DR GO; GO:0015668; F:Type III site-specific deoxyribonuclease ac. .; IEA.
DR GO; GO:0009307; P:DNA restriction; IEA.
DR InterPro; IPR006935; Resili.
DR Pfam; PF04851; Resili; 1.
SQ SEQUENCE 1098 AA; 121958 MW; 6470B478424986B CRC64;
Query Match 91.3%; Score 94; DB 2; Length 1098;
Best Local Similarity 78.3%; Pred. No. 2.5e-08;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 1 VEVXVDXFGREKDXIIISLSCV 23
DB 800 VEASVDAFGREKDXIIISLSCV 822
RESULT 13
ID Q7ZVZ4 PRELIMINARY; PRT; 1100 AA.
AC Q7ZVZ4;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Regulator of nonsense transcripts 1.
GN ORFNames=zgc:55472;

```

```

OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045353; AAH45353.1;
DR ZFIN; ZDB-GENE-040426-2835; zgc:55472.
DR GO; GO:0015668; F:Type III site-specific deoxyribonuclease ac. .; IEA.
DR GO; GO:0009307; P:DNA restriction; IEA.
DR InterPro; IPR006935; Resili.
DR Pfam; PF04851; Resili; 1.
SQ SEQUENCE 1100 AA; 122062 MW; 3215FEBDC6353CB CRC64;
Query Match 91.3%; Score 94; DB 2; Length 1100;
Best Local Similarity 78.3%; Pred. No. 2.5e-08;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 1 VEVXVDXFGREKDXIIISLSCV 23
DB 801 VEASVDAFGREKDXIIISLSCV 823
RESULT 14
ID RNT1_MOUSE STANDARD; PRT; 1113 AA.
AC Q9EPD0; Q9EPD0; Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Regulator of nonsense transcripts 1 (Nonsense mRNA reducing factor 1)
GN (NORF1) (Up-frameshift suppressor 1 homolog)
GN Name=Rent1; Synonym=hupfl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=21067878; PubMed=11152657; DOI=10.1093/hmg/10.2.99;
RA Medghalchi S.M., Friesmeyer P.A., Mendell J.T., Kelly A.G.,
RA Lawler A.M., Dietz H.C.;
RT "Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential
RT for mammalian embryonic viability."

```

RL Hum. Mol. Genet. 10:99-105(2001).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RA Seig M., Strande J., Beck-Engesser G.B.J., Liehr T., Winkler T.,  
 RA Jack H.-M.;  
 RT "Genomic structure, chromosomal localization and expression of murine  
 RT nonsense mRNA reducing factor 1 (MORF1)."  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carrinci P., Prange C.,  
 RA Raba S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.M., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 73-1102 FROM N.A.  
 RC STRAIN=129/Sv;  
 RX MEDLINE=97008109; PubMed=8855285; DOI=10.1073/pnas.93.20.10928;  
 RA Perllick H.A., Medghalchi S.M., Spencer F.A., Kendzior R.J. Jr.,  
 RA Dietz H.C.;  
 RT "Mammalian orthologues of a yeast regulator of nonsense transcript  
 RT stability."  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:10928-10932(1996).  
 CC -1- FUNCTION: Eliminates the production of nonsense-containing RNAs.  
 CC Essential for embryonic viability.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF322655; AAG4830.1; -;  
 CC EMBL; BC052149; AAH52149.1; -;  
 CC EMBL; AF182947; AAK08652.1; -;  
 CC MGD; MGI:107995; Renti.  
 CC InterPro; IPR006935; Resili.  
 CC ATP-binding; Helicase; Hydrolase; Zinc-finger.  
 KM DOMAIN 47 75 Ala/Gly/Pro-rich.  
 FT DOMAIN 1026 1113 Gln/Ser-rich.  
 FT ZN\_FING 126 154 C2H2-type (atypical) (Potential).  
 FT ZN\_FING 178 208 C4-type (Potential).  
 FT NP\_BIND 487 494 ATP (Potential).  
 FT CONFLICT 60 60 A -> P (in Ref. 2).  
 SQ SEQUENCE 1113 AA; 122656 MW; BF8BD50AD49DA54 CRC64;  
 Query Match 91.3%; Score 94; DB 1; Length 1113;  
 Best Local Similarity 78.3%; Pred. No. 2.5e-08;  
 Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 816 VEASVDAFOGREKDXIILSCVR 838  
 ||: || ||||| |||||  
 RESULT 15  
 ID Q6GYPS PRELIMINARY; PRT, 1113 AA.  
 AC Q6GYPS  
 DT 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE RENT1.  
 GN Name=Rent1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SvJ;  
 RA Seig M., Strande J.L., Liehr T., Roth B., Beck-Engesser G.B.,  
 RA Winkler T.H., Jack H.-M.;  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY597038; AAT46119.1; JOINED.  
 DR EMBL; AY597039; AAT46119.1; -;  
 DR InterPro; IPR006935; Resili.  
 DR Pfam; PF04851; Resili.1.  
 SQ SEQUENCE 1113 AA; 122787 MW; 689B4472FE93B576 CRC64;  
 Query Match 91.3%; Score 94; DB 2; Length 1113;  
 Best Local Similarity 78.3%; Pred. No. 2.5e-08;  
 Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEAVTVDXFOGREKDXIILSCVR 23  
 ||: || ||||| |||||  
 Search completed: April 18, 2005, 08:15:48  
 Job time : 97.0292 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: April 18, 2005, 07:10:15 ; Search time 89.6058 Seconds  
(without alignments)  
77.692 Million cell updates/sec

Title: US-10-652-334-9

Sequence: 1 IGFLXDKRINVALTRAK 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database : A.GeneSeq.16Dec04:\*

1: \_geneseqp1980s:.\*  
2: \_geneseqp1980s:.\*  
3: \_geneseqp2000s:.\*  
4: \_geneseqp2001s:.\*  
5: \_geneseqp2002s:.\*  
6: \_geneseqp2003as:.\*  
7: \_geneseqp2003bs:.\*  
8: \_geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	75	94.9	18 3	AAV77812 Motif IX
2	75	94.9	18 7	ABW01185 Saccharom
3	75	94.9	18 8	ADP44111 Yeast tra
4	73	92.4	415 3	AAV77813 Yeast Mtc
5	73	92.4	415 7	ABW01201 Saccharom
6	73	92.4	415 8	ADP44127 Yeast hel
7	73	92.4	992 8	ADS23757 Bacterial
8	73	92.4	1944 8	ADN19868 Bacterial
9	72	91.1	98 5	ABP07670 Human ORF
10	71	89.9	471 3	AAV77815 Yeast Sen
11	71	89.9	472 7	ABW01202 Saccharom
12	71	89.9	472 8	ADP44128 Yeast hel
13	71	89.9	2000 6	ABR53451 Protein s
14	71	89.9	2000 7	ADK64622 Disease t
15	71	89.9	2231 8	ADN19177 Bacterial
16	70	88.6	648 8	ADS41904 Bacterial
17	70	88.6	650 8	ADN20231 Bacterial
18	70	88.6	653 8	ADS43215 Bacterial
19	70	88.6	654 4	AAAB9566 Putative
20	70	88.6	655 4	AAAB62031 Recombina
21	70	88.6	656 8	ADN18615 Bacterial
22	69	87.3	373 8	ADN10398 Human pro
23	69	87.3	688 8	ADN47941 Thermoco
24	69	87.3	797 2	AAW36508 Human REN
25	69	87.3	925 8	ADN19434 Bacterial

26	69	87.3	935 7	ADJ69900 Human hea
27	69	87.3	1043 2	AAW36509 Murine RE
28	69	87.3	1118 6	ABG73900 Human REN
29	69	87.3	1140 2	AAV21377 Human HUP
30	68	86.1	380 3	AAV77814 Yeast Upf
31	68	86.1	380 7	ABW01205 Saccharom
32	68	86.1	380 8	ADP44131 Yeast hel
33	68	86.1	971 3	AAV98057 Yeast Upf
34	68	86.1	971 6	ABR53412 Protein s
35	68	86.1	971 7	ADK64706 Disease t
36	67	84.8	611 8	ADS21255 Bacterial
37	67	84.8	642 8	ADS43096 Bacterial
38	67	84.8	648 8	ADS42927 Bacterial
39	67	84.8	663 8	ADS43005 Bacterial
40	65	82.3	171 3	AAAB33009 Pinus rad
41	65	82.3	988 8	AAO26745 988-mer r
42	64	81.0	693 8	ADN22619 Bacterial
43	64	79.7	818 8	ADK68058 Female re
44	62	78.5	698 7	ADN25464 Hyperther
45	62	78.5	712 8	ADS44332 Bacterial

## ALIGNMENTS

RESULT 1	AAV77812	standard; peptide; 18 AA.
ID	AAV77812	
AC	AAV77812	
DT	31-MAY-2000	(first entry)
XX		
DE	Motif IX comprised in a gene modulating translation termination.	
KW	Helicase B; HCSB; MTT1; modulator of translation termination; eRF1; eRF3; eukaryotic release factor; peptidyl transferase; beta-thalassemia;	
KW	beta-globin; Duchene/Becker Muscular Dystrophy; antiannemic.	
XX		
OS	Unidentified.	
FX		
FT	Key	Location/Qualifiers
FT	Misc-difference 1. .18	/note= "residues indicated Xaa are unspecified"
XX		
XX	WO200005586-A2.	
XX	03-FEB-2000.	
XX	22-JUL-1999;	99WO-US016802.
XX	22-JUL-1998;	98US-00120435.
XX	(UYNE-) UNIV NEW JERSEY.	
XX		
PI	Peltz S, Czaplinski K, Dimman JD;	
XX		
DR	WPI; 2000-171458/15.	
XX		
PT	New multiprotein complex which can modulate peptidyl transferase activity	
PT	during translation, useful to treat diseases associated with peptidyl	
PT	transferase activity e.g. Duchene/Becker Muscular Dystrophy.	
XX		
PS	Clam 41; Page 80; 89pp; English.	
XX		
CC	The invention provides a new multiprotein complex which can modulate	
CC	peptidyl transferase activity during translation. The complex comprises	
CC	the gene encoding Helicase B (HCSB; renamed MTT1, for Modulator of	
CC	Translation Termination) and the conserved proteins known to interact and	
CC	carry out translation termination in eukaryotic cells, peptidyl	
CC	eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to	
CC	modulate peptidyl transferase activity during translation in a cell. It	
CC	can be administered therapeutically combined with a carrier in	

pharmaceutical compositions to treat diseases associated with peptidyl transferase activity, especially diseases resulting from a nonsense or frameshift mutation e.g. beta-thalassemia, beta-globin, Duchenne/Becker Muscular Dystrophy etc. It can be used to identify disease conditions involving a defect in the complex, by transfecting cells with encoding nucleic acid and determining the proportion of defective complex before and after transfection. It is also useful to screen for drugs involved in peptidyl transferase activity during translation, inhibiting the interaction between MT11 and eRF3 or involved in enhancing the translation termination. Vectors comprising polynucleotides encoding the complex (or antisense sequences) can be constructed and introduced into cells to interfere with complex expression and/or modulate the efficiency of translation termination of mRNA and/or degradation of aberrant transcripts in a cell. Agents binding to the complex can be identified and included in therapeutic compositions useful as above, and/or used to modulate peptidyl transferase activity during translation in cells. They are also useful to modulate the efficiency of translation termination of mRNA at a nonsense codon and/or promote degradation of aberrant transcripts in cells. The method can be used to identify agents/compositions modulating binding to MT11, useful to identify genes. Sequences AA77804-812 represent motifs I-IX comprised in the genes of interest, used for modulating translation termination

Sequence 18 AA;

Query Match 94.9%; Score 75; DB 3; Length 18;  
Best Local Similarity 100.0%; Pred. No. 2,4e-07; Mismatches 0; Gaps 0;  
Matches 18; Conservative 0; Indels 0;

Qy 1 IGFLDXRRINVALTRAK 18  
Db 1 IGFLDXRRINVALTRAK 18

RESULT 2  
ABW01185  
ID ABW01185 standard; peptide; 18 AA.

XX AC ABW01185;  
XX DT 15-JAN-2004 (first entry)

XX Saccharomyces cerevisiae motif IX peptide.

XX KW Modulator of translation termination; MT11; helicase B; antiviral;  
XX KM therapy; HCSB; nonsense mutation; yeast.

XX OS Saccharomyces cerevisiae.

XX FH Key Location/Qualifiers

XX FT Misc-difference 5 /label= Unknown

XX FT Misc-difference 7 /note= "Xaa may be any amino acid"

XX FT Misc-difference 7 /label= Unknown

XX FT Misc-difference 7 /note= "Xaa may be any amino acid"

XX FN US6630294-B1.

XX PD 07-OCT-2003.

XX PF 22-JUL-1999; 99US-00359268.

XX PR 22-JUL-1998; 98US-0093685P.

XX PA (UYNE-) UNIV NEW JERSEY MEDICINE &amp; DENTISTRY.

XX PI Peltz S, Czaplinski K, Dinman JD;

XX DR WPI; 2003-810549/76.

XX PT Identifying an agent that increases nonsense suppression, for antiviral  
PT therapy, by contacting modulator of translation termination (Mtl1) in

PT Saccharomyces cerevisiae with a test agent, and detecting specific  
PT binding to Mtl1.

XX PS Disclosure; Col 45-46; Opp; English.

XX CC The invention relates to a method of identifying an agent that increases  
CC nonsense suppression, by contacting modulator of translation termination  
CC (MT11) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.

XX CC The method is useful for identifying compositions or agents which  
CC increase nonsense suppression. The invention may also be used for  
CC antiviral therapy and for suppression of pathological nonsense mutations.  
XX CC The present sequence is Saccharomyces cerevisiae motif IX peptide

Sequence 18 AA;

Query Match 94.9%; Score 75; DB 7; Length 18;  
Best Local Similarity 100.0%; Pred. No. 2,4e-07; Mismatches 0; Gaps 0;  
Matches 18; Conservative 0; Indels 0;

Qy 1 IGFLDXRRINVALTRAK 18  
Db 1 IGFLDXRRINVALTRAK 18

RESULT 3  
ADP44111  
ID ADP44111 standard; peptide; 18 AA.

XX AC ADP44111;

XX DT 18-NOV-2004 (first entry)

XX DE Yeast translation termination modulation protein motif IX.

XX KM gene therapy; translation termination; RNA helicase; MT11;  
XX KM frameshift frequency; aberrant transcript degradation;  
XX KM peptidyl transferase modulation; beta-thalassemia; beta-globin;  
XX KM Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
XX KM Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;  
XX KM Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic fibrosis;  
XX KM Kidney Stone; Familial hypercholesterolaemia; Retinitis Pigmentosa;  
XX KM Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; Yeast.

XX OS Saccharomyces cerevisiae.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Misc-difference 5 /note= "Any amino acid"

XX FT Misc-difference 7 /note= "Any amino acid"

XX FT Misc-difference 7 /note= "Any amino acid"

XX FN US2004115787-A1.

XX PD 17-JUN-2004.

XX PF 28-AUG-2003; 2003US-00652334.

XX PR 22-JUL-1998; 98US-0093685P.

XX PR 22-JUL-1999; 99US-00359268.

XX PA (PELTZ/) PELTZ S.

XX PI Peltz S, Czaplinski K, Dinman JD;

XX DR WPI; 2004-449400/42.

XX PT Identifying a test composition or agent that modulates the efficiency of  
PT translation termination comprises contacting the MT11 with the test  
PT composition or agent, and determining if the test composition or agent  
PT inhibits the MT11.

XX Claim 41; SEQ ID NO 9; 41bp; English.

PS The invention relates to a method of identifying a test composition that

CC modules the efficiency of translation termination comprising contacting

CC the RNA helicase MTT1 with a composition or agent under conditions

CC permitting binding between the MTT1 and the composition, detecting

CC specific binding of the test composition or agent to the MTT1, and

CC determining if the test composition or agent inhibits the MTT1. The

CC composition and methods are useful for modulating the fidelity of

CC translation termination or for identifying agents that affect the

CC functional activity of mRNAs by altering frameshift frequency, permit

CC monitoring of a termination event, promote degradation of aberrant

CC transcripts, and provide modulators (inhibitor/stimulators) of peptidyl

CC transferase activity during initiation, elongation, termination and mRNA

CC degradation of translation. The agents, which may be antagonists or

CC agonists, are useful in screening, diagnostic and therapeutic purposes,

CC for diseases or conditions resulting from or cause premature translation,

CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular

CC Dystrophy, Hemophilia A, Haemophilia B, Von Willebrand Disease,

CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,

CC Hirschsprung disease, Cystic fibrosis, kidney stones, Familial

CC hypercholesterolemia, Retinitis Pigmentosa, or Neurofibromatosis,

CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents

CC the amino acid sequence of the yeast translation termination modulation

CC protein motif IX.

XX Sequence 18 AA;

SO Query Match 94.9%; Score 75; DB 8; Length 18;

Best Local Similarity 100.0%; Pred. No. 2.4e-07;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IGPLDXRRINVAALTRAK 18

DB 1 IGPLDXRRINVAALTRAK 18

RESULT 4

AAV77813

ID AAV77813 standard; peptide; 415 AA.

XX AAV77813;

AC AAV77813;

XX 31-MAY-2000 (first entry)

DT Yeast Mtt1 protein fragment.

XX Helicase B; HCSB; MTT1; modulator of translation termination; eRF1; eRF3;

KM eukaryotic release factor; peptidyl transferase; beta-thalassemia;

KW beta-globin; Duchenne/Becker Muscular Dystrophy; anti-anemic; yeast;

KM helicase.

XX Saccharomyces cerevisiae.

OS WO200005586-A2.

PN 03-FEB-2000.

PD 22-JUL-1999; 99WO-US016802.

PF 22-JUL-1998; 98US-00120435.

PR 22-JUL-1998; 98US-00120435.

XX (UYNE-) UNIV NEW JERSEY.

PA Peltz S, Czaplinski K, Dimman JD;

XX WPI; 2000-171458/15.

DR New multiprotein complex which can modulate peptidyl transferase activity

PT during translation, useful to treat diseases associated with peptidyl

PT transferase activity e.g. Duchenne/Becker Muscular Dystrophy.

XX

PS Example 1; Fig 1; 89bp; English.

XX The invention provides a new multiprotein complex which can modulate

CC peptidyl transferase activity during translation. The complex comprises

CC the gene encoding Helicase B (HCSB; renamed MTT1, for Modulator of

CC translation Termination) and the conserved proteins known to interact and

CC carry out translation termination in eukaryotic cells, peptidyl

CC eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to

CC modulate peptidyl transferase activity during translation in a cell. It

CC can be administered therapeutically combined with a carrier in

CC pharmaceutical compositions to treat diseases associated with peptidyl

CC transferase activity, especially diseases resulting from a nonsense or

CC frameshift mutation e.g. beta-thalassemia, beta-globin, Duchenne/Becker

CC Muscular Dystrophy etc. It can be used to identify disease conditions

CC involving a defect in the complex, by transfecting cells with encoding

CC nucleic acid and determining the proportion of defective complex before

CC and after transfection. It is also useful to screen for drugs involved in

CC peptidyl transferase activity during translation, inhibiting the

CC interaction between MTT1 and eRF3 or involved in enhancing translation

CC termination. Vectors comprising polynucleotides encoding the complex (or

CC antisense sequences) can be constructed and introduced into cells to

CC interfere with complex expression and so modulate the efficiency of

CC translation termination of mRNA and/or degradation of aberrant

CC transcripts in a cell. Agents binding to the complex can be identified

CC and included in therapeutic compositions useful as above, and/or used to

CC modulate peptidyl transferase activity during translation in cells. They

CC are also useful to modulate the efficiency of translation termination of

CC mRNA at a nonsense codon and/or promote degradation of aberrant

CC transcripts in cells. The method can be used to identify agents/

CC compositions modulating binding to MTT1, useful to identify genes.

CC Sequences AAV77813-817 represent protein fragments from yeast superfamily

CC group I helicases

XX Sequence 415 AA;

SO Query Match 92.4%; Score 73; DB 3; Length 415;

Best Local Similarity 83.3%; Pred. No. 1.9e-05;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 IGPLDXRRINVAALTRAK 18

DB 398 IGPLDXRRINVAALTRAK 415

RESULT 5

ABW01201

ID ABW01201 standard; protein; 415 AA.

XX ABW01201;

AC ABW01201;

XX 15-JAN-2004 (first entry)

DT Saccharomyces cerevisiae modulator of translation termination protein.

XX Modulator of translation termination; MTT1; helicase B; antiviral;

KM therapy; HCSB; nonsense mutation; yeast.

KW Saccharomyces cerevisiae.

OS US6630294-B1.

PN 07-OCT-2003.

PD 22-JUL-1999; 99US-00359268.

PF 22-JUL-1998; 98US-0093685P.

PR (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

PA Peltz S, Czaplinski K, Dimman JD;

XX WPI; 2003-810549/76.

DR

XX

PT Identifying an agent that increases nonsense suppression, for antiviral  
PT therapy, by contacting modulator of translation termination (Mtt1) in  
PT Saccharomyces cerevisiae with a test agent, and detecting specific  
PT binding to Mtt1.  
XX  
XX Disclosure; Col 49-52; 0pp; English.  
XX  
CC The invention relates to a method of identifying an agent that increases  
CC nonsense suppression, by contacting modulator of translation termination  
CC (MTT1) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.  
CC The method is useful for identifying compositions or agents which  
CC increase nonsense suppression. The invention may also be used for  
CC antiviral therapy and for suppression of pathological nonsense mutations.  
CC The present sequence is Saccharomyces cerevisiae MTT1 protein  
XX  
SQ Sequence 415 AA;  
Query Match 92.4%; Score 73; DB 7; Length 415;  
Best Local Similarity 83.3%; Pred. No. 1.9e-05;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 IGFIXDXRRINVALTRAK 18  
Db 398 IGFILDRKRRINVALTRAK 415  
RESULT 6  
ADP44127  
ID ADP44127 standard; protein; 415 AA.  
XX  
AC ADP44127;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Yeast helicase Mtt1.  
XX  
XX gene therapy; translation termination; RNA helicase; MTT1;  
XX frameshift frequency; aberrant transcript degradation;  
XX peptidyl transferase modulation; beta-thalassemia; beta-globin;  
XX Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
XX Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;  
XX Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic fibrosis;  
XX Kidney Stone; Familial hypercholesterolemia; Retinitis Pigmentosa;  
XX Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; yeast; enzyme.  
OS Saccharomyces cerevisiae.  
XX  
XX US2004115787-A1.  
XX  
XX 17-JUN-2004.  
XX  
XX 28-AUG-2003; 2003US-00652334.  
XX  
XX 22-JUL-1998; 98US-0093685P.  
XX  
XX 22-JUL-1999; 99US-00359268.  
XX  
XX (PELTZ/) PELTZ S.  
XX (CZAP/) CZAPINSKI K.  
XX (DINM/) DINMAN J D.  
XX  
XX Pelcz S, Czaplinski K, Dinman JD;  
XX WPI; 2004-449400/42.  
XX  
XX Identifying a test composition or agent that modulates the efficiency of  
XX translation termination comprises contacting the MTT1 with the test  
XX composition or agent, and determining if the test composition or agent  
XX inhibits the MTT1.  
XX  
XX Disclosure; SEQ ID NO 25; 41pp; English.  
XX  
CC The invention relates to a method of identifying a test composition that  
CC modulates the efficiency of translation termination comprising contacting

CC the RNA helicase MTT1 with a composition or agent under conditions  
CC permitting binding between the MTT1 and the composition, detecting  
CC specific binding of the test composition or agent to the MTT1, and  
CC determining if the test composition or agent inhibits the MTT1. The  
CC composition and methods are useful for modulating the fidelity of  
CC translation termination or for identifying agents that affect the  
CC functional activity of mRNAs by altering frameshift frequency, permit  
CC monitoring of a termination event, promote degradation of aberrant  
CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl  
CC transferase activity during initiation, elongation, termination and mRNA  
CC degradation of translation. The agents, which may be antagonists or  
CC agonists, are useful in screening, diagnostic and therapeutic purposes,  
CC for diseases or conditions resulting from or cause premature translation,  
CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular  
CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,  
CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,  
CC Hirschsprung disease, Cystic fibrosis, Kidney Stones, Familial  
CC hypercholesterolemia, Retinitis Pigmentosa, or Neurofibromatosis,  
CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents  
CC the amino acid sequence of the yeast helicase Mtt1.  
XX  
SQ Sequence 415 AA;  
Query Match 92.4%; Score 73; DB 8; Length 415;  
Best Local Similarity 83.3%; Pred. No. 1.9e-05;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 IGFIXDXRRINVALTRAK 18  
Db 398 IGFILDRKRRINVALTRAK 415  
RESULT 7  
ADS23757  
ID ADS23757 standard; protein; 992 AA.  
XX  
XX ADS23757;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
XX Bacterial polypeptide #12790.  
XX  
XX Recombinant DNA construct; transformed plant; improved plant property;  
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
XX pathogen tolerance; pest tolerance; plant disease resistance;  
XX cell cycle pathway modification; plant growth regulator;  
XX homologous recombination; seed oil yield; protein yield; carbohydrate;  
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
XX bacterial polypeptide.  
XX  
XX Bacteria.  
XX  
XX US2003233675-A1.  
XX  
XX 18-DEC-2003.  
XX  
XX 20-FEB-2003; 2003US-00369493.  
XX  
XX 21-FEB-2002; 2002US-0360039P.  
XX  
XX (CAOY/) CAO Y.  
XX (HINK/) HINKLE G J.  
XX (SLAT/) SLATER S C.  
XX (CHEN/) CHEN X.  
XX (GOLD/) GOLDMAN B S.  
XX  
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX WPI; 2004-061375/06.  
XX  
XX New recombinant DNA construct comprising a promoter positioned to provide  
XX for expression of a polynucleotide encoding a polypeptide from a  
XX microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 12790; 122pp; English.

PS The invention relates to a recombinant DNA construct comprising a

XX promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a

CC microbial source. The invention also relates to a transformed plant

CC comprising the recombinant DNA construct and a method of producing a

CC transformed plant having an improved property. The plant is a crop plant

CC such as maize or soybean. The method of producing a transformed plant

CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the

CC polynucleotide or polypeptide is useful for improving plant properties.

CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of

CC homologous recombination, modified seed oil or protein yield and/or

CC content, improved yield by modification of carbohydrate, nitrogen or

CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress

CC condition, improved lignin production or improved galactomanan

CC production. This sequence represents a bacterial polypeptide used in the

CC scope of the invention. Note: The sequence data for this patent did not

CC form part of the printed specification but was obtained in electronic

CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX

XX Sequence 992 AA;

SQ

Query Match 92.4%; Score 73; DB 8; Length 992;

Best Local Similarity 83.3%; Pred. No. 5.2e-05;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGPLDKRRINVALTRAK 18

|||||:|||||

Db 718 IGPLSDPRRLNVALTRAK 735

RESULT 8

ADN19868

ID ADN19868 standard; protein; 1944 AA.

XX

AC ADN19868;

XX

DT 02-DEC-2004 (first entry)

XX

DE Bacterial polypeptide #2521.

XX

KM Recombinant DNA construct; transformed plant; improved plant property;

KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

KM pathogen tolerance; pest tolerance; plant disease resistance;

KM cell cycle pathway modification; plant growth regulator;

KM homologous recombination; seed oil yield; protein yield; carbohydrate;

KM nitrogen; phosphorus; photosynthesis; lignin; galactomanan;

KM bacterial polypeptide.

XX

OS Bacteria.

XX

PN US200323675-A1.

XX

PD 18-DEC-2003.

XX

PP 20-FEB-2003; 2003US-00369493.

XX

PR 21-FEB-2002; 2002US-0360039P.

XX

XX (CAOY/) CAO Y.

PA (HINKLE/) HINKLE G J.

PA (SLATER/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX

XX WPI; 2004-061375/06.

DR

XX

XX New recombinant DNA construct comprising a promoter positioned to provide

PT for expression of a polynucleotide encoding a polypeptide from a

PT microbial source, useful for producing plants with improved properties.

XX

PS Claim 1; SEQ ID NO 2521; 122pp; English.

XX

CC The invention relates to a recombinant DNA construct comprising a

CC promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a

CC microbial source. The invention also relates to a transformed plant

CC comprising the recombinant DNA construct and a method of producing a

CC transformed plant having an improved property. The plant is a crop plant

CC such as maize or soybean. The method of producing a transformed plant

CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the

CC polynucleotide or polypeptide is useful for improving plant properties.

CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of

CC homologous recombination, modified seed oil or protein yield and/or

CC content, improved yield by modification of carbohydrate, nitrogen or

CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress

CC condition, improved lignin production or improved galactomanan

CC production. This sequence represents a bacterial polypeptide used in the

CC scope of the invention. Note: The sequence data for this patent did not

CC form part of the printed specification but was obtained in electronic

CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX

XX Sequence 1944 AA;

SQ

Query Match 92.4%; Score 73; DB 8; Length 1944;

Best Local Similarity 83.3%; Pred. No. 0.00011;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGPLDKRRINVALTRAK 18

|||||:|||||

Db 1729 IGPLQDLRLNVALTRAK 1746

RESULT 9

ABP07670

ID ABP07670 standard; protein; 98 AA.

XX

AC ABP07670;

XX

DT 25-JUN-2002 (first entry)

XX

DE Human ORFX protein sequence SEQ ID NO:15322.

XX

KM Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;

KM hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;

KM degenerative disorder; osteoarthritis; neurodegenerative disorder;

KM cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;

KM hypertension; hypothyroidism; cholesterol ester storage disease;

KM immune deficiency; immune disorder; infectious disease;

KM autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;

KM myasthenia gravis.

XX

OS Homo sapiens.

XX

PN WO200192523-A2.

XX

PD 06-DEC-2001.

XX

PP 29-MAY-2001; 2001WO-US010836.

XX

	PA	(CUBA-) CUBAGEN CORP.	
	PB		
	PR	Shimkets RA, Leach MD;	
	PS	WPI; 2002-106308/14.	
	PT	N-PSSD; ABN234422.	
	PP		
	PP	Novel human polypeptides and polynucleotides useful for diagnosing,	
	PP	preventing and treating cardiovascular disease, neurodegenerative,	
	PP	hyperproliferative disorders and autoimmune disorders.	
	PS	Disclosure; SEQ ID NO 15322; 1037pp; English.	
	XX		
	CC	The present invention describes substantially purified human proteins	
	CC	(referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1	
	CC	in the specification). ABN15762 to ABN27252 encode the human ORFX	
	CC	protein given in ABP00010 to ABP11500. ORFX proteins are useful for	
	CC	treating or preventing a pathology associated with an ORFX-associated	
	CC	disorder in humans, and in the manufacture of a medicament for treating a	
	CC	syndrome associated with ORFX-associated disorder. ORFX polynucleotide	
	CC	sequences can be used in gene therapy. ORFX sequences can be used in the	
	CC	treatment of cancer, hyperproliferative disorders, cirrhosis of liver,	
	CC	psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,	
	CC	osteoarthritis, neurodegenerative disorders, disorders related to organ	
	CC	transplantation, cardiovascular diseases, diabetes mellitus, systemic	
	CC	lupus erythematosus, hypertension, hypothyroidism, cholesterol ester	
	CC	storage disease, various immune deficiencies and disorders, infectious	
	CC	diseases, autoimmune disorders such as multiple sclerosis, rheumatoid	
	CC	arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host	
	CC	disease and autoimmune inflammatory eye disease. ORFX proteins are also	
	CC	useful for treating burns, incisions, ulcers, for treating osteoporosis,	
	CC	bone degenerative disorders, or periodontal disease, and for gut	
	CC	protection or regeneration and treatment of lung or liver fibrosis,	
	CC	reflexion injury in various tissues and conditions resulting from	
	CC	systemic cytokine damage. N.B. The sequence data for this patent did not	
	CC	form part of the printed specification, but was obtained in electronic	
	CC	format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences	
	SQ	Sequence 98 AA:	
OY	Query Match	91.1%; Score 72; DB 5; Length 98;	
ID	Best Local Similarity	77.8%; Pred. No. 6e-06; Indels 0; Gaps 0;	
XX	Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;		
OY	1 IGFLDXRRINVALTRAK 18		
ID	::   ::     :		
DB	71 IGFSDSRRLNVALTRAR 88		
RESULT 10			
ID	AAAY77815		
XX	AAAY77815 standard; peptide; 471 AA.		
AC			
XX	AAAY77815;		
DY			
DT	31-MAY-2000 (first entry)		
DE			
XX	Yeast SenI protein fragment.		
KM	Helicase B; HCSB; MTM1; modulator of translation termination; eRF1, eRF3;		
KM	eukaryotic release factor; peptidyl transferase; beta-thalassemia;		
KM	beta-globin; Duchene/Becker Muscular Dystrophy; antiemic; yeast;		
XX	helicase; SenI.		
OS	Saccharomyces cerevisiae.		
FN	WO200005586-A2.		
DD	03-FEB-2000.		

PF 22-JUL-1999; 99WO-US016802.  
XX  
PR 22-JUL-1998; 98US-00120435.  
XX  
PA (UTNE-) UNIV NEW JERSEY.  
XX  
PI Peltz S, Czaplinski K, Dimman JD;  
XX  
DR WPI, 2000-171458/15.  
XX  
PT New multiprotein complex which can modulate peptidyl transferase activity  
XX during translation, useful to treat diseases associated with peptidyl  
XX transferase activity e.g. Duchenne/Becker Muscular Dystrophy.  
XX  
PS Example 1; Fig 1; 89pp; English.  
XX  
CC The invention provides a new multiprotein complex which can modulate  
XX peptidyl transferase activity during translation. The complex comprises  
XX the gene encoding Helicase B (HCSB; renamed MTI1, for Modulator of  
XX Translation Termination) and the conserved proteins known to interact and  
XX carry out translation termination in eukaryotic cells, peptidyl  
XX eukaryotic release factor (eRF) 1 and eRF3. The complex can be used to  
XX modulate peptidyl transferase activity during translation in a cell. It  
XX can be administered therapeutically combined with a carrier in  
XX pharmaceutical compositions to treat diseases associated with peptidyl  
XX transferase activity, especially diseases resulting from a nonsense or  
XX frameshift mutation e.g. beta-thalassemia, beta-globin, Duchenne/Becker  
XX Muscular Dystrophy etc. It can be used to identify disease conditions  
XX involving a defect in the complex, by transfecting cells with encoding  
XX nucleic acid and determining the proportion of defective complex before  
XX and after transfection. It is also useful to screen for drugs involved in  
XX peptidyl transferase activity during translation, inhibiting the  
XX interaction between MTI1 and eRF3 or involved in enhancing translation  
XX termination. Vectors comprising polynucleotides encoding the complex (or  
XX antisense sequences) can be constructed and introduced into cells to  
XX interfere with complex expression and so modulate the efficiency of  
XX translation termination of mRNA and/or degradation of aberrant  
XX transcripts in a cell. Agents binding to the complex can be identified  
XX and included in therapeutic compositions useful as above, and/or used to  
XX modulate peptidyl transferase activity during translation in cells. They  
XX are also useful to modulate the efficiency of translation termination of  
XX mRNA at a nonsense codon and/or promote degradation of aberrant  
XX transcripts in cells. The method can be used to identify agents/  
XX compositions modulating binding to MTI1, useful to identify genes.  
XX Sequences AA177813-817 represent protein fragments from yeast superfamily  
XX group I helicases  
XX  
SQ Sequence 471 AA;

Query Match 89.9%; Score 71; DB 3; Length 471;  
Best local Similarity 77.8%; Pred. No. 5.5e-05;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0

CY 1 IGFLDXKRINVAALTRAK 18  
: ||| : ||| ||||  
Db 454 VGFLKDFRNMVVALTRAK 471

RESULT 11  
ABM01202  
ID ABM01202 standard; protein; 472 AA.

AC ABM01202;  
DT 15-JAN-2004 (first entry)  
DE Saccharomyces cerevisiae SEN1 protein.  
KW Modulator of translation termination; MTI1; helicase B; antiviral;  
XX therapy; HCSB; nonsense mutation; yeast.  
XX  
XX Saccharomyces cerevisiae.

```

PN US6630294-B1.
XX
PD 07-OCT-2003.
XX
PF 22-JUL-1999; 99US-00359268.
XX
PR 22-JUL-1998; 98US-0093685P.
XX
PA (UNNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
PI Peltz S, Czaplinski K, Dinman JD;
XX
DR WPI; 2003-810549/76.
XX
PT Identifying an agent that increases nonsense suppression, for antiviral
PT therapy, by contacting modulator of translation termination (MTT1) in
PT Saccharomyces cerevisiae with a test agent, and detecting specific
PT binding to Mtt1.
XX
PS Disclosure; Col 51-54; 0pp; English.
XX
CC The invention relates to a method of identifying an agent that increases
CC nonsense suppression, by contacting modulator of translation termination
CC (MTT1) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.
CC The method is useful for identifying compositions or agents which
CC increase nonsense suppression. The invention may also be used for
CC antiviral therapy and for suppression of pathological nonsense mutations.
CC The present sequence is Saccharomyces cerevisiae SEN1 protein
XX
SQ Sequence 472 AA;
XX
Query Match 89.9%; Score 71; DB 7; Length 472;
Best Local Similarity 77.8%; Pred. No. 5.5e-05;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 IGPLDKRRINVALTRAK 18
DB 455 VGFLKDFRBMNVVALTRAK 472
XX
RESULT 12
ADP44128
ID ADP44128 standard; protein; 472 AA.
XX
AC ADP44128;
XX
DT 18-NOV-2004 (first entry)
XX
DE Yeast helicase Sen1.
XX
KW gene therapy; translation termination; RNA helicase; MTT1;
KW frameshift frequency; aberrant transcript degradation;
KW peptidyl transferase modulation; beta-thalassemia; beta-globin;
KW Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;
KW Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;
KW Ovarian Cancer; Wilms Tumour; Hirschprung disease; Cystic fibrosis;
KW Kidney Stone; Familial hypercholesterolaemia; Retinitis Pigmentosa;
KW Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; yeast; enzyme.
XX
OS Saccharomyces cerevisiae.
XX
PN US2004115787-A1.
XX
PD 17-JUN-2004.
XX
PF 28-AUG-2003; 2003US-00652334.
XX
PR 22-JUL-1998; 98US-0093685P.
XX
PR 22-JUL-1999; 99US-00359268.
XX
PA (PELTZ S,
PA (CZAPLINSKI K,
PA (DINMAN J D.

```

```

XX
PI Peltz S, Czaplinski K, Dinman JD;
XX
DR WPI; 2004-449400/42.
XX
PT Identifying a test composition or agent that modulates the efficiency of
PT translation termination comprises contacting the MTT1 with the test
PT composition or agent, and determining if the test composition or agent
PT inhibits the MTT1.
XX
PS Disclosure; SEQ ID NO 26; 41pp; English.
XX
CC The invention relates to a method of identifying a test composition that
CC modulates the efficiency of translation termination comprising contacting
CC the RNA helicase MTT1 with a composition or agent under conditions
CC permitting binding between the MTT1 and the composition, detecting
CC specific binding of the test composition or agent to the MTT1, and
CC determining if the test composition or agent inhibits the MTT1. The
CC composition and methods are useful for modulating the fidelity of
CC translation termination or for identifying agents that affect the
CC functional activity of mRNAs by altering frameshift frequency, permit
CC monitoring of a termination event, promote degradation of aberrant
CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl
CC transferase activity during initiation, elongation, termination and mRNA
CC degradation of translation. The agents, which may be antagonists or
CC agonists, are useful in screening, diagnostic and therapeutic purposes,
CC for diseases or conditions resulting from or cause premature translation,
CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular
CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,
CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,
CC Hirschprung disease, Cystic fibrosis, Kidney Stones, Familial
CC hypercholesterolaemia, Retinitis Pigmentosa, or Neurofibromatosis,
CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents
CC the amino acid sequence of the yeast helicase Sen1.
XX
SQ Sequence 472 AA;
XX
Query Match 89.9%; Score 71; DB 8; Length 472;
Best Local Similarity 77.8%; Pred. No. 5.5e-05;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 IGPLDKRRINVALTRAK 18
DB 455 VGFLKDFRBMNVVALTRAK 472
XX
RESULT 13
ABR53451
ID ABR53451 standard; protein; 2000 AA.
XX
AC ABR53451;
XX
DT 20-JUN-2003 (first entry)
XX
DE Protein sequence #SEQ ID 1767.
XX
KW Multiprotein complex; eukaryote; drug target; diagnosis.
XX
OS Saccharomyces cerevisiae.
XX
PN EP1258494-A1.
XX
PD 20-NOV-2002.
XX
PF 20-DEC-2001; 2001EP-00130253.
XX
PR 15-MAY-2001; 2001EP-00111774.
XX
PA (CELL-) CELLZOME AG.
XX
PI Bauer A, Gavin A, Grandt P, Krause R, Krause UD, Kuester BD,
PI Marzloch M, Schultz JD, Superti-Furga GD;
XX

```



DR WPI: 2003-250078/25.  
DR N-PSDB; ACC61493.  
XX  
PT New isolated protein complexes useful for diagnosing a disease or  
PT disorder, or as a target for an active agent of a pharmaceutical,  
PT preferably a drug target in the treatment or prevention of disease or  
PT disorder.  
XX  
PS Disclosure; SEQ ID NO 1767, 17pp + Sequence Listing; English.  
XX  
CC The invention relates to multiprotein complexes from eukaryotes. Proteins  
CC of the invention and DNA sequences encoding them are given in records  
CC AB552568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are  
CC obtainable by using a protein as a bait and isolating the set of proteins  
CC which is attached thereto from cells. Such protein complexes may comprise  
CC up to 30 distinct proteins. Protein complexes of the invention are useful  
CC for diagnosing a disease or disorder, or as a target for an active agent  
CC of a pharmaceutical, preferably a drug target in the treatment or  
CC prevention of a disease or disorder. Note: The sequence data for this  
CC patent is not represented in the printed specification, but is based on  
CC sequence information supplied by the European Patent Office. The complete  
CC document is available on CD-ROM  
XX  
SQ Sequence 2000 AA;  
XX  
Query Match 89.9%; Score 71; DB 6; Length 2000;  
Best Local Similarity 77.8%; Pred. No. 0.00028;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
XX  
QY 1 IGFLLDXRRINVALTRAK 18  
:|||||:|||||  
Db 1805 VGFLKDFRPMNVALTRAK 1822  
XX  
RESULT 14  
ADK64622  
ID ADK64622 standard; protein; 2000 AA.  
XX  
AC ADK64622;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Disease treating protein complex-derived protein #1064.  
XX  
KM protein complex; drug target; diagnosis.  
XX  
OS Unidentified.  
XX  
PN EP1338608-A2.  
XX  
PD 27-AUG-2003.  
XX  
PF 20-DEC-2002; 2002EP-00102902.  
XX  
PR 20-DEC-2001; 2001EP-00130253.  
XX  
PA (CELL-) CELLZOME AG.  
XX  
PI Bauer A, Gavin A, Superti-Furga G, Kueser B, Schultz J,  
PI Marzloch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A,  
PI Michon A, Leutwein C, Rick J;  
XX  
DR WPI: 2003-638460/61.  
DR N-PSDB; ADK64623.  
XX  
XX  
PT New proteins and protein complexes from eukaryotes, useful as targets in  
PT drug screening, or in diagnosing or screening for the presence of a  
PT disease or disorder, or a predisposition for developing a disease or  
PT disorder in a subject.  
XX  
PS Disclosure; SEQ ID NO 2127, 13pp; English.  
XX  
CC The invention relates to novel protein complexes comprising a first and a

CC second protein, or its derivative, fragment, homologue or variant. The  
CC proteins are selected from given protein complexes, which are not defined  
CC in the specification. The variants are encoded by nucleic acids that  
CC hybridize to the nucleic acids encoding the proteins under low stringency  
CC conditions. The protein complexes are useful as targets for an active  
CC agent of a pharmaceutical. These protein complexes are particularly  
CC useful as drug targets for the treatment or preventing of a disease or  
CC disorder. The complexes and methods above are useful in diagnosing or  
CC screening for the presence of a disease or disorder or a predisposition  
CC for developing a disease or disorder in a subject. These are also useful  
CC in screening for a drug for treatment or prevention of a disease or  
CC disorder. The molecule that modulates the amount, activity or protein  
CC components of the complex is useful for the manufacture of a medicament  
CC for the treatment or prevention of a disease or disorder. This sequence  
CC corresponds to a protein of the invention. (Note: the sequence data for  
CC this patent did not form part of the printed specification but was  
CC obtained from the EPO in electronic format).  
XX  
SQ Sequence 2000 AA;  
XX  
Query Match 89.9%; Score 71; DB 7; Length 2000;  
Best Local Similarity 77.8%; Pred. No. 0.00028;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
XX  
QY 1 IGFLLDXRRINVALTRAK 18  
:|||||:|||||  
Db 1805 VGFLKDFRPMNVALTRAK 1822  
XX  
RESULT 15  
ADN19177  
ID ADN19177 standard; protein; 2231 AA.  
XX  
AC ADN19177;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Bacterial polypeptide #1830.  
XX  
KM Recombinant DNA construct; transformed plant; improved plant property;  
KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;  
KM pathogen tolerance; pest tolerance; plant disease resistance;  
KM cell cycle pathway modification; plant growth regulator;  
KM homologous recombination; seed oil yield; protein yield; carbohydrate;  
KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KM bacterial polypeptide.  
XX  
OS Bacteria.  
XX  
PN US2003233675-A1.  
XX  
PD 18-DEC-2003.  
XX  
PF 20-FEB-2003; 2003US-00369493.  
XX  
PR 21-FEB-2002; 2002US-0360039P.  
XX  
PA (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX  
DR WPI: 2004-061375/06.  
XX  
PT New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
PS Claim 1; SEQ ID NO 1830, 122pp; English.  
XX



CC The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition, improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX  
 SQ Sequence 2231 AA;

Query Match 89.9%; Score 71; DB 8; Length 2231;  
 Best Local Similarity 77.8%; Pred. No. 0.00031;  
 Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLDXRINVALTRAK 18  
 :||| |:|||||  
 Db 1805 VGFLKDFRMNVALTRAK 1822

Search completed: April 18, 2005, 08:03:51  
 Job time : 89.717 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 18, 2005, 07:37:21 ; Search time 22.4672 Seconds  
(without alignments)  
59.807 Million cell updates/sec

Title: US-10-652-334-9

Sequence: 1 IGFLXDXRRINVALTRAK 18

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

Issued Patents AA: \*  
1: /cgn2\_6/ptcdat1/1/aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptcdat1/1/aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptcdat1/1/aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptcdat1/1/aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptcdat1/1/aa/6C.COMB.pep: \*  
6: /cgn2\_6/ptcdat1/1/aa/backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	94.9	18	US-09-359-268A-9	Sequence 9, Appl1
2	73	92.4	415	US-09-359-268A-25	Sequence 25, Appl1
3	73	92.4	917	US-09-248-796A-19347	Sequence 19347, A
4	72	91.1	405	US-09-248-796A-19107	Sequence 19107, A
5	71	89.9	472	US-09-359-268A-26	Sequence 26, Appl1
6	69	87.3	1043	US-08-724-354D-4	Sequence 4, Appl1
7	69	87.3	1043	US-09-270-984A-4	Sequence 4, Appl1
8	69	87.3	1118	US-08-724-354D-2	Sequence 2, Appl1
9	69	87.3	1118	US-09-270-984A-2	Sequence 2, Appl1
10	69	87.3	1140	US-09-949-016-10116	Sequence 10116, A
11	68	86.1	211	US-09-902-540-16148	Sequence 16148, A
12	68	86.1	380	US-09-359-268A-29	Sequence 29, Appl1
13	68	86.1	971	US-08-724-354D-22	Sequence 22, Appl1
14	68	86.1	971	US-09-270-984A-22	Sequence 22, Appl1
15	68	86.1	971	US-09-177-431-8	Sequence 8, Appl1
16	65	82.3	171	US-09-640-211A-1058	Sequence 1058, Ap
17	65	82.3	993	US-09-538-092-1100	Sequence 1100, Ap
18	63	79.7	426	US-09-248-796A-15170	Sequence 15170, A
19	58	73.4	414	US-09-359-268A-28	Sequence 28, Appl1
20	58	73.4	683	US-09-538-092-483	Sequence 483, App
21	57	72.2	3177	US-08-477-451-4	Sequence 4, Appl1
22	52	65.8	366	US-09-359-268A-27	Sequence 27, Appl1
23	49	62.0	11	US-09-359-268A-24	Sequence 24, Appl1
24	48	60.8	157	US-09-270-767-32463	Sequence 32463, A
25	48	60.8	157	US-09-270-767-47680	Sequence 47680, A
26	48	60.8	444	US-09-270-767-44429	Sequence 44429, A
27	47	59.5	181	US-09-270-767-31838	Sequence 31838, A

28	47	59.5	181	US-09-270-767-47055	Sequence 47055, A
29	46	58.2	219	US-09-248-796A-18933	Sequence 18933, A
30	46	58.2	238	US-09-270-767-56745	Sequence 56745, A
31	46	58.2	486	US-09-270-767-41521	Sequence 41521, A
32	44	55.7	2108	US-09-252-991A-31502	Sequence 31502, A
33	41	51.9	345	US-09-538-092-588	Sequence 588, App
34	40	50.6	513	US-09-902-540-13564	Sequence 13564, A
35	39	49.4	426	US-09-902-540-12700	Sequence 12700, A
36	39	49.4	1213	US-09-543-681A-6478	Sequence 6478, Ap
37	38	48.1	253	US-09-489-039A-10712	Sequence 10712, A
38	38	48.1	1233	US-09-194-613-5	Sequence 5, Appl1
39	37	46.8	169	US-09-270-767-46729	Sequence 46729, A
40	37	46.8	177	US-08-700-013B-11	Sequence 11, Appl1
41	37	46.8	177	US-08-700-013B-13	Sequence 13, Appl1
42	37	46.8	251	US-09-191-468-72	Sequence 72, Appl1
43	37	46.8	251	US-09-191-468-74	Sequence 74, Appl1
44	37	46.8	251	US-09-191-468-76	Sequence 76, Appl1
45	37	46.8	251	US-09-191-468-78	Sequence 78, Appl1

## ALIGNMENTS

```
RESULT 1
US-09-359-268A-9
; Sequence 9, Application US/09359268A
; Patent No. 6630294
; GENERAL INFORMATION:
; APPLICANT: Peltz, Stuart
; APPLICANT: Czaplinski, Kevin
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
; FILE REFERENCES: 601-1-85N
; CURRENT APPLICATION NUMBER: US/09/359,268A
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,685
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: saccharomyces cerevisiae
; FEATURES:
; OTHER INFORMATION: Xaa = any amino acid
US-09-359-268A-9

Query Match          94.9%; Score 75; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.5e+08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 IGFLXDXRRINVALTRAK 18
Db      1 IGFLXDXRRINVALTRAK 18

RESULT 2
US-09-359-268A-25
; Sequence 25, Application US/09359268A
; Patent No. 6630294
; GENERAL INFORMATION:
; APPLICANT: Peltz, Stuart
; APPLICANT: Czaplinski, Kevin
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES
; FILE REFERENCES: 601-1-85N
; CURRENT APPLICATION NUMBER: US/09/359,268A
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,685
```

;; PRIOR FILING DATE: 1998-07-22  
;; NUMBER OF SEQ ID NOS: 32  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 25  
;; LENGTH: 415  
;; TYPE: PRT  
;; ORGANISM: saccharomyces cerevisiae  
US-09-359-268A-25

Query Match 92.4%; Score 73; DB 4; Length 415;  
Best Local Similarity 83.3%; Pred. No. 3.1e-06;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLEXRRINVALTRAK 18  
Db 398 IGFLEXRRINVALTRAK 415

RESULT 3  
US-09-248-796A-19347  
; Sequence 19347, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 19347  
; LENGTH: 917  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-19347

Query Match 92.4%; Score 73; DB 4; Length 917;  
Best Local Similarity 83.3%; Pred. No. 7.8e-06;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLEXRRINVALTRAK 18  
Db 860 IGFLEXRRINVALTRAK 877

RESULT 4  
US-09-248-796A-19107  
; Sequence 19107, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 19107  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: Candida albicans  
; NAME/KEY: UNSURE  
; LOCATION: (1)  
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno

US-09-248-796A-19107

Query Match 91.1%; Score 72; DB 4; Length 405;  
Best Local Similarity 77.8%; Pred. No. 4.8e-06;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLEXRRINVALTRAK 18  
Db 142 VGFLSDVRRMVALTRAK 159

RESULT 5  
US-09-359-268A-26  
; Sequence 26, Application US/09359268A  
; Patent No. 6630294  
; GENERAL INFORMATION:  
; APPLICANT: Peltz, Stuart  
; APPLICANT: Czaplinski, Kevin  
; APPLICANT: Dimman, Jonathan D.  
; TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
; TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: 601-1-85N  
; CURRENT APPLICATION NUMBER: US/09/359,268A  
; CURRENT FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 60/093,685  
; PRIOR FILING DATE: 1998-07-22  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 26  
; LENGTH: 472  
; TYPE: PRT  
; ORGANISM: saccharomyces cerevisiae  
US-09-359-268A-26

Query Match 89.9%; Score 71; DB 4; Length 472;  
Best Local Similarity 77.8%; Pred. No. 9.1e-06;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLEXRRINVALTRAK 18  
Db 455 VGFLDXRRINVALTRAK 472

RESULT 6  
US-08-724-354D-4  
; Sequence 4, Application US/08724354D  
; Patent No. 5994119  
; GENERAL INFORMATION:  
; APPLICANT: Dietz, Harry C.  
; TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
; TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,354D  
; FILING DATE: 01-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/016,482  
; FILING DATE: 29-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halle, Lisa A.

REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1043 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-724-354D-4

Query Match 87.3%; Score 69; DB 2; Length 1043;  
Best Local Similarity 77.8%; Pred. No. 5.6e-05;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IGFLLDKRRINVALTRAK 18  
Db 773 IGFLLDPRLNVALTRAR 790

RESULT 7  
US-09-270-984A-4  
Sequence 4, Application US/09270984A  
Patent No. 6048965  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/270,984A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/724,354  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1043 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-270-984A-4

Query Match 87.3%; Score 69; DB 3; Length 1043;  
Best Local Similarity 77.8%; Pred. No. 5.6e-05;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IGFLLDKRRINVALTRAK 18  
Db 773 IGFLLDPRLNVALTRAR 790

RESULT 8  
US-08-724-354D-2  
Sequence 2, Application US/08724354D  
Patent No. 5994119  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,354D  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/016,482  
FILING DATE: 29-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1118 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-724-354D-2

Query Match 87.3%; Score 69; DB 2; Length 1118;  
Best Local Similarity 77.8%; Pred. No. 6e-05;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IGFLLDKRRINVALTRAK 18  
Db 850 IGFLLDPRLNVALTRAR 867

RESULT 9  
US-09-270-984A-2  
Sequence 2, Application US/09270984A  
Patent No. 6048965  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/270,984A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/724,354  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1118 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-270-984A-2

Query Match 87.3%; Score 69; DB 3; Length 1118;  
Best Local Similarity 77.8%; Pred. No. 6e-05;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLLDXRRINVAALTRAK 18  
DB 850 IGFLLDXRRINVAALTRAK 867

RESULT 10  
US-09-949-016-10116  
Sequence 10116, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYNOPEPTIDES IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CLO01307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10116  
LENGTH: 1140  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-10116

Query Match 87.3%; Score 69; DB 4; Length 1140;  
Best Local Similarity 77.8%; Pred. No. 6.2e-05;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLLDXRRINVAALTRAK 18  
DB 872 IGFLLDXRRINVAALTRAK 869

RESULT 11  
US-09-902-540-16148  
Sequence 16148, Application US/09902540  
Patent No. 6833447  
GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Wiegand, Roger C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

FILE REFERENCE: 38-10(15849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 16148  
LENGTH: 211  
TYPE: PRT  
ORGANISM: Myxococcus xanthus  
US-09-902-540-16148

Query Match 86.1%; Score 68; DB 4; Length 211;  
Best Local Similarity 72.2%; Pred. No. 1.4e-05;  
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLLDXRRINVAALTRAK 18  
DB 151 IGFLLDXRRINVAALTRAK 168

RESULT 12  
US-09-359-268A-29  
Sequence 29, Application US/09359268A  
Patent No. 6630294  
GENERAL INFORMATION:  
APPLICANT: Peltz, Stuart  
APPLICANT: Czaplinski, Kevin  
APPLICANT: Dimman, Jonathan D.  
TITLE OF INVENTION: A SUBFAMILY OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
FILE REFERENCE: 601-1-85N  
CURRENT APPLICATION NUMBER: US/09/359,268A  
CURRENT FILING DATE: 1999-07-22  
PRIOR APPLICATION NUMBER: 60/093,685  
PRIOR FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 29  
LENGTH: 380  
TYPE: PRT  
ORGANISM: saccharomyces cerevisiae  
US-09-359-268A-29

Query Match 86.1%; Score 68; DB 4; Length 380;  
Best Local Similarity 77.8%; Pred. No. 2.8e-05;  
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 IGFLLDXRRINVAALTRAK 18  
DB 363 IGFLLDXRRINVAALTRAK 380

RESULT 13  
US-08-724-354D-22  
Sequence 22, Application US/08724354D  
Patent No. 5994119  
GENERAL INFORMATION:  
APPLICANT: Dietz, Harry C.  
TITLE OF INVENTION: MAMMALIAN REGULATOR OF  
TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible

```
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,354D
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/016,482
FILING DATE: 29-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hallie, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 971 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-724-354D-22

Query Match      86.1%; Score 68; DB 2; Length 971;
Best Local Similarity 77.8%; Pred. No. 8.1e-05;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1  IGFLLDXRRINVALTRAK 18
DB      786  IGFLLDRPRRLNVGLTRAK 803

RESULT 14
US-09-270-984A-22
Sequence 22, Application US/09270984A
Patent No. 6048965
GENERAL INFORMATION:
APPLICANT: Dietz, Harry C.
TITLE OF INVENTION: MAMMALIAN REGULATOR OF
NUMBER OF INVENTION: NONSENSE-MEDIATED RNA DECAY
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/270,984A
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/724,354
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hallie, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 971 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-270-984A-22
```

```
Query Match      86.1%; Score 68; DB 3; Length 971;
Best Local Similarity 77.8%; Pred. No. 8.1e-05;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1  IGFLLDXRRINVALTRAK 18
DB      786  IGFLLDRPRRLNVGLTRAK 803

RESULT 15
US-09-177-431-8
Sequence 8, Application US/09177431
Patent No. 6071700
GENERAL INFORMATION:
APPLICANT: He, Feng
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE
NUMBER OF INVENTION: ABSENCE OF NONSENSE-MEDIATED RNA DECAY FUNCTION
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/177,431
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/955,472
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Passer, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07917/050001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-9806
TELFX: 200154
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 971 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-09-177-431-8

Query Match      86.1%; Score 68; DB 3; Length 971;
Best Local Similarity 77.8%; Pred. No. 8.1e-05;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1  IGFLLDXRRINVALTRAK 18
DB      786  IGFLLDRPRRLNVGLTRAK 803

Search completed: April 18, 2005, 08:18:37
Job time : 23.4672 secs
```

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 18, 2005, 08:06:16 ; Search time 63.854 Seconds  
(without alignments)  
93.693 Million cell updates/sec

Title: US-10-652-334-9

Sequence: 1 IGFLXDXRRINVALTRAK 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppaa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppaa/US10D\_PUBCOMB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*  
21: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*  
22: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	75	94.9	18	US-10-652-334-9	Sequence 9, Appl1
2	73	92.4	415	US-10-652-334-9	Sequence 25, Appl1
3	73	92.4	992	US-10-369-493-12790	Sequence 12790, A
4	73	92.4	1944	US-10-369-493-12790	Sequence 12790, A
5	71	89.9	472	US-10-652-334-9	Sequence 25, Appl1
6	71	89.9	2231	US-10-369-493-1830	Sequence 1830, Ap
7	70	88.6	123	US-10-424-599-21689	Sequence 21689, A
8	70	88.6	439	US-10-767-701-41078	Sequence 41078, A
9	70	88.6	648	US-10-369-493-20334	Sequence 20334, A
10	70	88.6	650	US-10-369-493-2884	Sequence 2884, Ap
11	70	88.6	653	US-10-369-493-21645	Sequence 21645, A
12	70	88.6	655	US-10-828-924-80	Sequence 80, Appl1
13	70	88.6	656	US-10-369-493-1268	Sequence 1268, Ap

14	69	87.3	415	US-10-425-114-45987	Sequence 45987, A
15	69	87.3	437	US-10-425-114-42857	Sequence 42857, A
16	69	87.3	449	US-10-425-114-57765	Sequence 57765, A
17	69	87.3	638	US-10-437-963-181690	Sequence 181690, A
18	69	87.3	734	US-10-424-599-205643	Sequence 205643, A
19	69	87.3	925	US-10-369-493-2087	Sequence 2087, Ap
20	69	87.3	935	US-10-408-765A-1706	Sequence 1706, Ap
21	69	87.3	1118	US-10-474-553-6	Sequence 6, Appl1
22	69	87.3	1361	US-10-437-963-165703	Sequence 165703, A
23	68	86.1	380	US-10-652-334-9	Sequence 29, Appl1
24	67	84.8	88	US-10-437-963-182848	Sequence 182848, A
25	67	84.8	161	US-10-425-114-38303	Sequence 38303, A
26	67	84.8	182	US-10-767-701-62775	Sequence 62775, A
27	67	84.8	250	US-10-424-599-236053	Sequence 236053, A
28	67	84.8	611	US-10-369-493-10288	Sequence 10288, A
29	67	84.8	642	US-10-369-493-21526	Sequence 21526, A
30	67	84.8	648	US-10-369-493-21526	Sequence 21526, A
31	67	84.8	663	US-10-369-493-21535	Sequence 21535, A
32	67	84.8	1975	US-10-437-963-140079	Sequence 140079, A
33	66	83.5	127	US-10-767-701-48804	Sequence 48804, A
34	66	83.5	235	US-10-425-114-42518	Sequence 42518, A
35	66	83.5	332	US-10-425-114-38637	Sequence 38637, A
36	66	83.5	350	US-10-424-599-241211	Sequence 241211, A
37	66	83.5	404	US-10-425-114-37833	Sequence 37833, A
38	66	83.5	626	US-10-437-963-166322	Sequence 166322, A
39	66	83.5	1027	US-10-437-963-185291	Sequence 185291, A
40	66	83.5	1323	US-10-437-963-111794	Sequence 111794, A
41	64	81.0	693	US-10-369-493-52712	Sequence 52712, Ap
42	63	79.7	317	US-10-425-114-42716	Sequence 42716, A
43	63	79.7	637	US-10-424-599-233501	Sequence 233501, A
44	63	79.7	639	US-10-425-114-37717	Sequence 37717, A
45	63	79.7	890	US-10-437-963-118530	Sequence 118530, A

## ALIGNMENTS

RESULT 1  
US-10-652-334-9  
Sequence 9, Application US/10652334  
Publication No. US20040115787A1  
GENERAL INFORMATION:  
APPLICANT: Pelcz, Stuart  
APPLICANT: Czaplinski, Kevin  
APPLICANT: Dimman, Jonathan D.  
TITLE OF INVENTION: A SUBSTRATE OF RNA HELICASES WHICH ARE MODULATORS OF  
TITLE OF INVENTION: THE FIDELITY OF TRANSLATION TERMINATION AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: 601-1-85N  
CURRENT FILING DATE: 2003-08-28  
PRIORITY APPLICATION NUMBER: US/09/359,268A  
PRIORITY FILING DATE: 1999-07-22  
PRIORITY APPLICATION NUMBER: 60/093,685  
PRIORITY FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 9  
LENGTH: 18  
TYPE: PRT  
ORGANISM: saccharomyces cerevisiae  
FEATURES:  
OTHER INFORMATION: Xaa = any amino acid  
US-10-652-334-9

Query Match 94.9%; Score 75; DB 16; Length 18;  
Best Local Similarity 100.0%; Pred. No. 6.6e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 IGFLXDXRRINVALTRAK 18  
Db 1 IGFLXDXRRINVALTRAK 18

RESULT 4  
US-10-369-493-2521  
; Sequence 2521, Application US/10369493  
; Publication No. US20030233675A1

;; APPLICANT: Cao, Yongwei  
;; APPLICANT: Hinkle, Gregory J.  
;; APPLICANT: Slater, Steven C.  
;; APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 1830  
LENGTH: 2231  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-10-369-493-1830

Query Match 89.9%; Score 71; DB 15; Length 2231;  
Best Local Similarity 77.8%; Pred. No. 0.0001;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLLDXRRINVAITRAK 18  
DB 1805 VGFLKDFRRNVAVITRAK 1822

RESULT 7  
US-10-424-599-216889  
Sequence 216889, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovallik David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 216889  
LENGTH: 123  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(123)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_3787C.1.pap  
US-10-424-599-216889

Query Match 88.6%; Score 70; DB 15; Length 123;  
Best Local Similarity 72.2%; Pred. No. 6e-06;  
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLLDXRRINVAITRAK 18  
DB 52 VGFLGDSRRINVAITRAK 69

RESULT 8  
US-10-767-701-41078  
Sequence 41078, Application US/10767701  
Publication No. US20040172684A1  
GENERAL INFORMATION:  
APPLICANT: Kovallik, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53535)B  
CURRENT APPLICATION NUMBER: US/10/767,701  
CURRENT FILING DATE: 2004-01-29

NUMBER OF SEQ ID NOS: 63128  
SEQ ID NO 41078  
LENGTH: 439  
TYPE: PRT  
ORGANISM: Sorghum bicolor  
FEATURE:  
OTHER INFORMATION: Clone ID: SORBI-28MAY03-C1946\_1.pap  
US-10-767-701-41078

Query Match 88.6%; Score 70; DB 16; Length 439;  
Best Local Similarity 72.2%; Pred. No. 2.6e-05;  
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLLDXRRINVAITRAK 18  
DB 366 VGFLGDSRRINVAITRAK 383

RESULT 9  
US-10-369-493-20334  
Sequence 20334, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 20334  
LENGTH: 648  
TYPE: PRT  
ORGANISM: Pyrococcus horikoshii  
US-10-369-493-20334

Query Match 88.6%; Score 70; DB 15; Length 648;  
Best Local Similarity 77.8%; Pred. No. 4e-05;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLLDXRRINVAITRAK 18  
DB 599 IGFLLDXRRINVAITRAK 616

RESULT 10  
US-10-369-493-2884  
Sequence 2884, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 2884  
LENGTH: 650  
TYPE: PRT  
ORGANISM: Thermotoga maritima

```
QY      1  IGLXDXRRINVALTRAK 18
        |||| | ||: ||: ||||
Db      599 IGLKDLRLNLSLTRAK 616
```

```
QY      1 IGFLDXKRINVALTRAK 18
      |||||:|||||:
Db      14 IGFLNDPRRLNVALTRAR 31
```

RESULT 15  
US-10-425-114-42857  
; Sequence 42857, Application US/10425114  
; Publication No. US20040034888A1

GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 42857  
LENGTH: 437  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: 700476463\_FLI.pep  
US-10-425-114-42857

Query Match 87.3%; Score 69; DB 15; Length 437;  
Best Local Similarity 77.8%; Pred. No. 4.1e-05;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLDXKRINVALTRAK 18  
DB 51 IGFLNDPRRLNVALTRAR 68

Search completed: April 18, 2005, 09:04:08  
Job time : 64.854 secs

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 18, 2005, 07:32:05 ; Search time 17.3431 Seconds  
(without alignments)  
99.861 Million cell updates/sec

Title: US-10-652-334-9

Sequence: 1 IGFLXDKRRINVALTRAK 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

## Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	92.4	1121	2	S30862 DNA dependent ATPa
2	73	92.4	1944	2	T40065 tRNA-splicing endo
3	72	91.1	692	2	E90113 hypothetical prote
4	71	89.9	1069	2	T43280 nonsense-mediated
5	71	89.9	2231	2	S53416 SEN1 protein - yea
6	70	88.6	245	2	D72258 helicase-related p
7	70	88.6	530	2	D70476 DNA helicase - Agu
8	70	88.6	650	2	G72429 hypothetical prote
9	70	88.6	653	2	B75105 probable DNA helic
10	70	88.6	656	2	B71080 probable DNA-bind
11	70	88.6	1687	2	T39072 DNA2-NAM7 helicase
12	70	88.6	2142	2	D86302 P17F16.1 protein -
13	69	87.3	935	2	S62476 hypothetical prote
14	69	87.3	1311	2	T08986 prematurely termi
15	68	86.1	971	2	S23408 transcription cont
16	67	84.8	642	2	D69085 DNA helicase homol
17	67	84.8	648	2	C69423 probable DNA helic
18	67	84.8	663	2	H64312 probable DNA-bind
19	66	83.5	1090	2	T00533 inulin II gene en
20	66	83.5	989	2	T48845 Ig mu chain switch
21	65	82.3	993	2	A47500 DNA-binding protei
22	65	82.3	993	2	S35633 hypothetical prote
23	64	81.0	693	2	T26415 related to SEN1 pr
24	64	81.0	1825	2	T53521 probable helicase
25	62	78.5	635	2	T02699 probable dna-bind
26	62	78.5	660	2	T41580 hypothetical prote
27	59	74.7	660	2	F85069 protein PIE2.16 l
28	59	74.7	1075	2	C9682 hypothetical prote
29	58	73.4	633	2	T28786

30	58	73.4	683	2	S34700 probable purine nu
31	57	72.2	274	2	D64588 probable DNA helic
32	57	72.2	1004	2	A39611 probable GTP-bind
33	57	72.2	1048	2	C86189 protein T25N20.11
34	56	70.9	239	2	T46441 hypothetical prote
35	56	70.9	678	2	T42668 hypothetical prote
36	56	70.9	1076	2	B96682 protein PIE2.14 l
37	55	69.6	1069	2	T22138 hypothetical prote
38	55	69.6	1105	2	T22132 hypothetical prote
39	54	68.4	1077	2	T50697 hypothetical prote
40	53	67.1	814	2	T00740 hypothetical prote
41	52	65.8	1522	2	S48904 probable purine nu
42	51	64.6	821	2	C84304 DNA helicase (Impo
43	50	63.3	2219	2	T27684 hypothetical prote
44	48	60.8	394	2	H81807 conserved hypothet
45	48	60.8	394	2	B81062 conserved hypothet

## ALIGNMENTS

## RESULT 1

S30862 DNA dependent ATPase/DNA helicase B - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YER176w

C:Species: Saccharomyces cerevisiae

C>Date: 28-May-1993 #sequence revision 28-May-1993 #text\_change 09-Jul-2004

C:Accession: S30862; S50679; J2C490; PC2368

R:Muligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Senl, P.; Komp, C.; Wei, Y.; Taylor,

submitted to the EMBL Data Library, February 1993

A:Reference number: S30812

A:Accession: S30862

A:Molecule type: DNA

A:Residues: 1-1121 <MUL>

A:Cross-references: UNIPROT:P32644; GB:U18922; EMBL:L11229; NID:9603405; PIDN:AA864703.1

R:Dietrich, F.S.

submitted to the EMBL Data Library, December 1994

A:Description: The sequence of S. cerevisiae cosmid 9163 and 9132.

A:Reference number: S50679

A:Accession: S50679

A:Molecule type: DNA

A:Residues: 1-1121 <DIR>

A:Cross-references: EMBL:U18922; NID:9603405; PIDN:AA864703.1; PID:9603417; MIPS:YER176w

R:Blasas, R.E.; Chen, P.H.; Leezyk, J.; Blasas, S.B.

Biochem. Biophys. Res. Commun. 206, 850-856, 1995

A:Title: Biochemical and genetic characterization of a replication protein A dependent D.

A:Reference number: J2C490; MUID:95134267; PMID:7832796

A:Accession: J2C490

A:Molecule type: DNA

A:Residues: 1-1121 <BIS>

A:Accession: PC2368

A:Molecule type: protein

A:Residues: 277-283;623-633; 'X',635-643 <BI2>

A:Comment: This enzyme plays pivotal roles in the unwinding of the DNA double helix dur

C:Genetics:

A:Gene: SGD:ECM32

A:Cross-references: SGD:S0000978; MIPS:YER176w

A:Map position: 5R

C:Keywords: nucleotide binding; P-loop

P/670-677/Region: nucleotide-binding motif A (P-loop)

Query Match 92.4%; Score 73; DB 2; Length 1121;

Best local similarity 83.3%; Pred. No. 1.4e-05;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Cy 1 IGFLXDKRRINVALTRAK 18

DB 1065 IGFLXDKRRINVALTRAK 1082

## RESULT 2

T40065

tRNA-splicing endonuclease positive effector - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: T40065  
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; LeJaure, V.; Gilbert, F.  
Submitted to the EMBL Data Library, December 1998  
A/Reference number: Z21903  
A/Accession: T40065  
A/Status: preliminary; translated from GB/EMBL/DDBJ  
A/Molecule type: DNA  
A/Residues: 1-1944 <LYN>  
A/Cross-references: UNIPROT:Q94387; EMBL:AL034463; PDB:CAA22438.1; GSPDB:GN00067; SPDB:  
C/Experimental source: Strain 972h-, cosmid c29A10  
C/Genetics:  
A/Gene: SPDB:SPBC29A10.10C  
A/Map position: 2

Query Match 92.4%; Score 73; DB 2; Length 1944;  
Best Local Similarity 83.3%; Pred. No. 2.4e-05;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

1 IGFLDXRRINVALTRAK 18  
1729 IGFLDXRRINVALTRAK 1746

RESULT 3  
E90113  
hypothetical protein component of a tRNA splicing complex [imported] - Guillardia theta  
C/Species: nucleomorph Guillardia theta  
A/Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C/Accession: E90113  
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rej  
Nature 410, 1091-1096, 2001  
A/Title: The highly reduced genome of an enslaved algal nucleus.  
A/Reference number: A99082; MUID:11323671; PMID:11323671  
A/Accession: E90113  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-692 <DOU>  
A/Cross-references: UNIPROT:Q9AVZ7; GB:AU010592; NID:g12580756; PDB:CAC27074.1; GSPDB:C  
C/Genetics:  
A/Gene: component of a tRNA splicing complex  
A/Map position: 2  
A/Genome: nucleomorph  
C/Keywords: nucleomorph

Query Match 91.1%; Score 72; DB 2; Length 692;  
Best Local Similarity 83.3%; Pred. No. 1.3e-05;  
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 IGFLDXRRINVALTRAK 18  
598 IGFLDXRRINVALTRAK 615

RESULT 4  
T43280  
nonsense-mediated mRNA decay trans-acting factor - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C/Accession: T43280  
R:Page, M.F.; Carr, B.; Anders, K.R.; Grimson, A.; Anderson, P.  
Mol. Cell. Biol. 19, 5943-5951, 1999  
A/Title: SMG-2 is a phosphorylated protein required for mRNA surveillance in Caenorhabdi  
A/Reference number: Z22389; MUID:99384262; PMID:10454541  
A/Accession: T43280  
A/Status: preliminary; translated from GB/EMBL/DDBJ  
A/Molecule type: mRNA  
A/Residues: 1-1069 <PAG>  
A/Cross-references: UNIPROT:O76512; EMBL:AF074017; NID:g3328176; PDB:AAC26789.1; PID:g3

Query Match 89.9%; Score 71; DB 2; Length 1069;

Best Local Similarity 77.8%; Pred. No. 3.2e-05;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

1 IGFLDXRRINVALTRAK 18  
824 IGFLDXRRINVALTRAK 841

RESULT 5  
S53416  
SEN1 protein - yeast (Saccharomyces cerevisiae)  
N/Alternate names: Protein L9576.1; protein YLR430w  
C/Species: Saccharomyces cerevisiae  
C/Date: 05-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 09-Jul-2004  
C/Accession: S53416; A44387; S41985  
R:Favell, A.  
submitted to the EMBL Data Library, February 1995  
A/Description: The sequence of S. cerevisiae cosmid 9576.  
A/Reference number: S53409  
A/Accession: S53416  
A/Molecule type: DNA  
A/Residues: 1-2231 <FAV>  
A/Cross-references: UNIPROT:Q00416; EMBL:U20939; NID:g664871; PID:g664872; MIPS:YLR430w  
A/Experimental source: Strain S288c (AB972)  
R:deMarini, D.J.; Winey, M.; Ursic, D.; Webb, F.; Culbertson, M.R.  
Mol. Cell. Biol. 12, 2154-2164, 1992  
A/Title: SEN1, a positive effector of tRNA-splicing endonuclease in Saccharomyces cerevis  
A/Reference number: A44387; MUID:9226590; PMID:1569945  
A/Accession: A44387  
A/Molecule type: DNA  
A/Residues: 'MS', 130, 'RCEREVO', 131-2231 <DEM>  
A/Cross-references: GB:M74589; NID:g172573; PDB:AAB63976.1; PID:g172574  
C/Genetics:  
A/Gene: SGD:SEN1  
A/Cross-references: SGD:S0004422; MIPS:YLR430w  
A/Map position: 12R  
C/Function:  
A/Description: may be component of nuclear splicing complex  
C/Keywords: nucleotide binding; nucleus; P-loop  
F1357-1364/Region: nucleotide-binding motif A (P-loop)

Query Match 89.9%; Score 71; DB 2; Length 2231;  
Best Local Similarity 77.8%; Pred. No. 6.8e-05;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

1 IGFLDXRRINVALTRAK 18  
1805 VGFLDXRRINVALTRAK 1822

RESULT 6  
D72258  
helicase-related protein - Thermotoga maritima (strain MSB8)  
C/Species: Thermotoga maritima  
C/Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C/Accession: D72258  
R:Nelson, K.B.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
C.M.  
Nature 399, 323-329, 1999  
A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
A/Reference number: A72200; MUID:99287316; PMID:10360571  
A/Accession: D72258  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-245 <ARN>  
A/Cross-references: UNIPROT:Q9X1D1; GB:AE001793; GB:AE000512; NID:g4981963; PDB:AAD3648;  
C/Experimental source: strain MSB8  
C/Genetics:  
A/Gene: TM411

Query Match 88.6%; Score 70; DB 2; Length 245;  
Best Local Similarity 77.8%; Pred. No. 1.1e-05;





Best Local Similarity 77.8%; Pred. No. 7.9e-05;  
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IGFIXDXRRINVALTRAK 18

Db 1591 IGFVADIRRMVALTRAK 1608

## RESULT 12

D86303 F1716.1 protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: D86303

R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Lin, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D86303

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-2142 <STO>

A:Cross-references: UNIPROT:Q9FWR3; GB:AE005172; NID:g9954728; PIDN:AG09081.1; GSPDB:GN

C:Genetics:

A:Map position: 1

Query Match 88.6%; Score 70; DB 2; Length 2142;

Best Local Similarity 77.8%; Pred. No. 0.0001;

Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IGFIXDXRRINVALTRAK 18

Db 1752 IGFVADIRRMVALTRAK 1769

## RESULT 13

S62476 hypothetical protein SPAC16C9.06C - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 16-May-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004

C:Accession: S62476; T37779

R:Badcock, K.; Churcher, C.M.

submitted to the EMBL Data Library, October 1995

A:Reference number: S62445

A:Accession: S62476

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-935 <BAD>

A:Cross-references: UNIPROT:Q09820; EMBL:Z54366; NID:g1019812; PIDN:CAA91194.1; PID:g132

R:Badcock, K.; Churcher, C.M.; Barrett, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, October 1995

A:Reference number: Z21745

A:Accession: T37779

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 11-935 <BA2>

A:Cross-references: EMBL:Z54366; PIDN:CAA91194.2; GSPDB:GN00066; SPDB:SPAC16C9.06C

A:Experimental source: strain 972h-; cosmid c16C9

C:Genetics:

A:Gene: SPDB:SPAC16C9.06C

A:Map position: 1L

Query Match 87.3%; Score 69; DB 2; Length 935;

Best Local Similarity 77.8%; Pred. No. 6.7e-05;

Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IGFIXDXRRINVALTRAK 18

Db 768 IGFVADIRRMVALTRAK 805

## RESULT 14

T08986 hypothetical protein F6G3.130 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C:Accession: T08986

R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X

submitted to the Protein Sequence Database, May 1999

A:Reference number: Z16520

A:Accession: T08986

A:Molecule type: DNA

A:Residues: 1-1311 <BEV>

A:Cross-references: UNIPROT:Q9SZW3; EMBL:AL078464; GSPDB:GN00062; ATSP:F6G3.130

A:Experimental source: cultivar Columbia; BAC clone F6G3

C:Genetics:

A:Gene: ATSP:F6G3.130

A:Map position: 4

A:Introns: 414/1; 474/1; 506/1; 547/2; 591/3; 1179/3

Query Match 87.3%; Score 69; DB 2; Length 1311;

Best Local Similarity 72.2%; Pred. No. 9.5e-05;

Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IGFIXDXRRINVALTRAK 18

Db 1158 IGFVADIRRMVALTRAK 1175

## RESULT 15

S23408 prematurely terminated mRNA decay factor NAM7 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YM9583.05c; protein YMR080c; UPF1 protein

C:Species: Saccharomyces cerevisiae

C>Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004

C:Accession: S23408; A44388; S54455

R:Altamira, N.; Groudinsky, O.; Dujardin, G.; Slonimski, P.P.

J. Mol. Biol. 224, 575-587, 1992

A:Title: NAM7 nuclear gene encodes a novel member of a family of helicases with a Zn-11g

A:Reference number: S23408; MUID:92235815; PMID:1314899

A:Accession: S23408

A:Molecule type: DNA

A:Residues: 1-971 <ALT>

A:Cross-references: UNIPROT:P30771; EMBL:X62394; NID:g4022; PIDN:CAA44266.1; PID:g4023

R:Leeds, P.; Wood, J.M.; Lee, B.S.; Culbertson, M.R.

Mol. Cell. Biol. 12, 2165-2177, 1992

A:Title: Gene products that promote mRNA turnover in Saccharomyces cerevisiae.

A:Reference number: A44388; MUID:92236591; PMID:1569946

A:Accession: A44388

A:Molecule type: DNA

A:Residues: 1-971 <LEB>

A:Cross-references: GB:476659; NID:g173141; PIDN:AAA35197.1; PID:g173142

R:Gentles, S.; Bowman, S.

submitted to the EMBL Data Library, May 1995

A:Reference number: S54451

A:Accession: S54451

A:Molecule type: DNA

A:Residues: 1-971 <GEN>

A:Cross-references: EMBL:Z49259; NID:g807956; PID:g807962; MIPS:YMR080C

A:Experimental source: strain AB972

C:Genetics:

A:Gene: SGD:NAM7; UPF1

A:Cross-references: SGD:S0004685; MIPS:YMR080C

A:Map position: 13R

F:430-437/Region: GTP binding; mitochondrial; nucleotide binding; nucleus, P-loop

F:545-548/Region: GTP-binding NKXD motif

Query Match 86.1%; Score 68; DB 2; Length 971;

Tue Apr 19 09:10:47 2005

US-10-652-334-9.FPR

Page 5

Best Local Similarity 77.8%; Pred. No. 0.00011;  
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

QY      1  IGFLXDKRRINVALTRAK  18
          |||||  ||:|||||
Db      786  IGFLRDPRLNVGLTRAK  803

```

Search completed: April 18, 2005, 08:06:05  
Job time : 17.3431 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 18, 2005, 07:33:55 ; Search time 75.1533 Seconds  
(without alignments)  
122.648 Million cell updates/sec

Title: US-10-652-334-9  
Perfect score: 79  
Sequence: 1 IGFLXDXRRINVALTRAK 18

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_03:.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	93.7	1124	2	06BG10 parametrium
2	73	92.4	964	2	06C803 yarrowia li
3	73	92.4	969	2	06C868 kluyveromyc
4	73	92.4	985	2	06BPM3 debaryomyc
5	73	92.4	1000	2	075DS7 ashyba goss
6	73	92.4	1060	2	08TFW3 aspergillus
7	73	92.4	1079	2	06MY12 aspergillus
8	73	92.4	1093	1	RNT1_NEURC
9	73	92.4	1121	1	YB06_YEAST
10	73	92.4	1124	2	06BNH2 debaryomyc
11	73	92.4	1125	2	06PKF3 candida gla
12	73	92.4	1139	2	07RKP6 plasmodium
13	73	92.4	1944	2	094387 schizosacch
14	73	92.4	2743	2	08IER9 plasmodium
15	73	92.4	692	2	09AVZ7 guillardia
16	73	92.4	1097	1	RNT1_FUGRU
17	72	91.1	1120	2	07PWZ4 anopheles g
18	72	91.1	1180	1	RNT1_DROME
19	72	91.1	1297	2	07ROL6 plasmodium
20	72	91.1	1554	2	08IJY4 plasmodium
21	71	89.9	583	2	Q7RAT3 baccharomyc
22	71	89.9	757	2	Q7LIE9 baccharomyc
23	71	89.9	1069	1	RNT1_CAEBL
24	71	89.9	1997	2	06CWA6 kluyveromyc
25	71	89.9	2027	2	0756Z8 ashyba goss
26	71	89.9	2231	1	SENI_YEAST
27	70	88.6	245	2	Q9XID1 thermocoga
28	70	88.6	527	2	06CHW3 yarrowia li
29	70	88.6	530	2	067840 aquifex ao
30	70	88.6	650	2	Q9WXM0 thermocoga
31	70	88.6	653	2	Q9UZB6 pyrococcus

32	70	88.6	655	2	Q8U398 pyrococcus
33	70	88.6	656	2	O58624 pyrococcus
34	70	88.6	763	2	Q7UMP1 rhodospirillum
35	70	88.6	1024	2	08IIT6 plasmodium
36	70	88.6	1687	1	SENI_SCHPO
37	70	88.6	2142	2	Q9FWR3
38	69	87.3	543	2	08K0N4
39	69	87.3	925	1	RNT1_SCHPO
40	69	87.3	1098	2	06GMR2
41	69	87.3	1100	2	07ZVZ4
42	69	87.3	1113	1	RNT1_MOUSE
43	69	87.3	1113	2	06GYP5
44	69	87.3	1118	2	086Z25
45	69	87.3	1124	2	06PHQ5
					Q8U398 pyrococcus
					O58624 pyrococcus
					Q7UMP1 rhodospirillum
					08IIT6 plasmodium
					092355 schizosacch
					Q9FWR3 arabidopsis
					08K0N4 mus musculus
					009820 schizosacch
					06GMR2 xenopus lae
					07ZVZ4 brachydanio
					Q8GPD mus musculus
					06GYP5 mus musculus
					086Z25 homo sapien
					06PHQ5 mus musculus

## ALIGNMENTS

RESULT 1	ID	Q6BG10	PRELIMINARY;	PRT;	1124 AA.
AC	Q6BG10				
DT	25-OCT-2004	(TREMBLrel. 28, Created)			
DT	25-OCT-2004	(TREMBLrel. 28, Last sequence update)			
DT	25-OCT-2004	(TREMBLrel. 28, Last annotation update)			
DE	TRNA-splicing endonuclease positive effector, putative.				
GN	ORFNames=PTM3.43c;				
OS	Parametrium tetraurelia.				
OC	Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Pentaculida;				
OX	Parametrium.				
NCBI	TaxID=5888;				
BN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Stock d4-2;				
RX	PubMed=15296759; DOI=10.1016/j.cub.2004.07.029;				
RA	Zagulski M., Nowak J.K., Le Mouel A., Nowacki M., Migdalski A.,				
RA	Gromadka R., Noel B., Blanc I., Dessen P., Winkler P., Keller A.M.,				
RA	Cohen J., Meyer B., Sperling L.;				
RT	"High Coding Density on the Largest Parametrium tetraurelia Somatic				
RT	Chromosome."				
RL	Curr. Biol. 14:1397-1404(2004).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Stock d4-2;				
RA	Nowak J.K., Migdalski A., Gromadka R., Zagulski M.;				
RT	"Parametrium megabase sequencing project."				
RL	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; CR548612; CA103240.1;				
DR	GO; GO:0004519; F:endonuclease activity; IEA.				
KW	Endonuclease.				
SO	SEQUENCE 1124 AA; 131764 MW; ECC6FED1566CB4B CRC64;				
Qy	Query Match	93.7%;	Score 74;	DB 2;	Length 1124;
	Best Local Similarity	83.3%;	Pred. No. 4.9e-05;		
Matches	15;	Conservative 1;	Mismatches 2;	Indels 0;	Gaps 0;
Db	1 IGFLXDXRRINVALTRAK 18				
	934 IGFLXDXRRINVALTRAK 951				
RESULT 2					
ID	Q6C803	PRELIMINARY;	PRT;	964 AA.	
AC	Q6C803				
DT	25-OCT-2004	(TREMBLrel. 28, Created)			
DT	25-OCT-2004	(TREMBLrel. 28, Last sequence update)			
DT	25-OCT-2004	(TREMBLrel. 28, Last annotation update)			
DE	Yarrowia lipolytica chromosome D of strain CL1899 of Yarrowia				
DE	lipolytica.				
GN	ORFNames=YALI0D23881g;				
OS	Yarrowia lipolytica CL1899.				

```

OC Burkayota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN
  [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla B.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne B., Bleykasten C.,
RA Boissière A., Boyer J., Cattoilco L., Confantolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantreaye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolaki M., Oztas S., Ozler-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.,
RA "Genome evolution in yeasts.",
RT Nature 430:35-44(2004).
RL
  [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope; (JUL-2004) to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382130; CAG81409.1; -.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00487; DEXDC; 1.
KW ATP-binding.
SQ SEQUENCE 964 AA; 107764 MW; 68C05A712597B8DD CRC64;

Query Match          92.4%; Score 73; DB 2; Length 964;
Best Local Similarity 83.3%; Pred. No. 6.6e-05;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Cy 1 IGFLLDXRRINVALTRAK 18
    |||||
Db 802 IGFLLDXRRINVALTRAK 819

RESULT 3
O6CWM68 PRELIMINARY; PRT; 969 AA.
AC O6CWM68;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Kluyveromyces lactis strain NRRL Y-1140 chromosome B of strain NRRL Y-
DE 1140 of Kluyveromyces lactis.
GN ORFNames=KLA0B06435g;
OS Kluyveromyces lactis NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;
RN
  [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla B.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne B., Bleykasten C.,
RA Boissière A., Boyer J., Cattoilco L., Confantolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantreaye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolaki M., Oztas S., Ozler-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.,
RA "Genome evolution in yeasts.",
RT Nature 430:35-44(2004).
RL
  [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RA Genoscope; (JUL-2004) to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382137; CAG8086.1; -.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR SMART; SM00487; DEXDC; 1.
SQ SEQUENCE 965 AA; 110705 MW; 14BDD331AE37B0D CRC64;

Query Match          92.4%; Score 73; DB 2; Length 965;
Best Local Similarity 83.3%; Pred. No. 6.6e-05;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.,
RA "Genome evolution in yeasts.",
RT Nature 430:35-44(2004).
RL
  [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RA Genoscope; (JUL-2004) to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382122; CAH0214.1; -.
DR InterPro; IPR006935; Resili.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF04851; Resili; 1.
DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN 1.
SQ SEQUENCE 969 AA; 108943 MW; 206C14F91B6BCE29 CRC64;

Query Match          92.4%; Score 73; DB 2; Length 969;
Best Local Similarity 83.3%; Pred. No. 6.6e-05;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Cy 1 IGFLLDXRRINVALTRAK 18
    |||||
Db 784 IGFLLDXRRINVALTRAK 801

RESULT 4
O6BPM3 PRELIMINARY; PRT; 985 AA.
AC O6BPM3;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Debaryomyces hanseni chromosome B of strain CBS767 of Debaryomyces
DE hanseni.
GN ORFNames=DEHA0E13002g;
OS Debaryomyces hanseni CBS767.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=284592;
RN
  [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla B.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne B., Bleykasten C.,
RA Boissière A., Boyer J., Cattoilco L., Confantolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantreaye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolaki M., Oztas S., Ozler-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.,
RA "Genome evolution in yeasts.",
RT Nature 430:35-44(2004).
RL
  [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RA Genoscope; (JUL-2004) to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382137; CAG8086.1; -.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR SMART; SM00487; DEXDC; 1.
SQ SEQUENCE 985 AA; 110705 MW; 14BDD331AE37B0D CRC64;

Query Match          92.4%; Score 73; DB 2; Length 985;
Best Local Similarity 83.3%; Pred. No. 6.6e-05;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

Best Local Similarity 83.3%; Pred. No. 6.7e-05;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLEXRRINVALTRAK 18  
DB 813 IGFLEXRRINVALTRAK 830

## RESULT 5

Q75DS7 PRELIMINARY; PRT; 1000 AA.

AC Q75DS7;  
DT 05-JUN-2004 (TREMBlrel. 27, Created)  
DT 05-JUN-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUN-2004 (TREMBlrel. 27, Last annotation update)  
DE ABR0222C.  
GN ORFNames=ABR022C;  
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.  
RN NCBI\_TaxID=33169;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 10895;  
RA Voegel S.E., Dietrich F.S., Brachat S., Iersch A., Gaffney T.,  
RA Philippsen P.;  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AEO16815; AAS50792.2;  
DR GO; GO:0000166; F:nucleotide binding; IEA.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR SMART; SM00382; AAA; 1.  
KM ATP-binding.  
SQ SEQUENCE 1000 AA; 111963 MW; 39DA257F675B72A6 CRC64;

Query Match 92.4%; Score 73; DB 2; Length 1000;  
Best Local Similarity 83.3%; Pred. No. 6.8e-05;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLEXRRINVALTRAK 18  
DB 795 IGFLEXRRINVALTRAK 812

## RESULT 6

Q8TFW3 PRELIMINARY; PRT; 1060 AA.

AC Q8TFW3;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Possible regulator of nonsense transcripts.  
GN Name=AfAsC11.22c;  
OS Aspergillus fumigatus (Sartorya fumigata).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
OX NCBI\_TaxID=5085;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Af293;  
RA Knowles D.G., Warren T., Hall N., Quail M., Woodward J.R.,  
RA Denning D.W., Anderson M.T., Barrall B.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL713629; CAD28448.1;  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0000166; F:nucleotide binding; IEA.  
DR GO; GO:0015668; F:type III site-specific deoxyribonuclease ac. .; IEA.  
DR InterPro; IPR003593; P:DNA restriction; IEA.  
DR InterPro; IPR011545; DEAD/DEAH\_N.  
DR InterPro; IPR006935; ResIII.  
DR Pfam; PF04851; ResIII; 1.  
DR SMART; SM00382; AAA; 1.  
DR SMART; SM00467; DEXDC; 1.

KM ATP-binding;  
SQ SEQUENCE 1060 AA; 116728 MW; 744DFC58A26EB77B CRC64;

Query Match 92.4%; Score 73; DB 2; Length 1060;  
Best Local Similarity 83.3%; Pred. No. 7.2e-05;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLEXRRINVALTRAK 18  
DB 786 IGFLEXRRINVALTRAK 803

## RESULT 7

Q6MYI2 PRELIMINARY; PRT; 1079 AA.

AC Q6MYI2;  
DT 05-JUN-2004 (TREMBlrel. 27, Created)  
DT 05-JUN-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUN-2004 (TREMBlrel. 27, Last annotation update)  
DE Regulator of nonsense transcripts, putative.  
GN ORFNames=AfAsC11.22c;  
OS Aspergillus fumigatus (Sartorya fumigata).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
RN NCBI\_TaxID=5085;  
RX [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=14998527; DOI=10.1016/j.fgb.2003.12.003;  
RX Pain A., Woodward J., Quail M.A., Anderson M.T., Clark R., Collins M.,  
RX Foster N., Fraser A., Harris D., Latke N., Murphy L., Humphray S.,  
RX O'Neill S., Petes M., Price C., Rabinowitch E., Rajandream M.A.,  
RX Salzberg S., Saunders D., Seegar K., Sharp S., Warren T.,  
RA Denning D.W., Barrall B., Hall N.;  
RT "Insight into the genome of Aspergillus fumigatus: analysis of a 922  
kb region encompassing the nitrate assimilation gene cluster.";  
RL Fungal Genet. Biol. 41:443-453(2004).  
DR EMBL; BX649606; CAF32021.1;  
DR InterPro; IPR011545; DEAD/DEAH\_N.  
DR InterPro; IPR006935; ResIII.  
DR Pfam; PF04851; ResIII; 1.  
DR SMART; SM00467; DEXDC; 1.  
SQ SEQUENCE 1079 AA; 118655 MW; A8907E6FA8500E63 CRC64;

Query Match 92.4%; Score 73; DB 2; Length 1079;  
Best Local Similarity 83.3%; Pred. No. 7.3e-05;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLEXRRINVALTRAK 18  
DB 805 IGFLEXRRINVALTRAK 822

## RESULT 8

RNT1\_NEUCR STANDARD; PRT; 1093 AA.

AC Q9HEH1; Q7RVU9;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 05-JUN-2004 (Rel. 44, Last annotation update)  
DE Regulator of nonsense transcripts 1 homolog.  
GN ORFNames=2E4.130; NCU04242.1;  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetales; Sordariiales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=74-OR23-1A / FGSC 987;  
RX MEDLINE=22542210; PubMed=12655011; DOI=10.1093/nar/gkg293;  
RA Manhaupt G., Montone C., Haase D., Wewes H.-W., Algen V.,  
RA Hohnel J.D., Fartmann B., Nyakatura G., Kempken F., Maier J.,  
RA Schulte U.;  
RT "What's in the genome of a filamentous fungus? Analysis of the

RT Neurospora genome sequence."  
 RL Nucleic Acids Res. 31:1944-1954(2003).  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=74-OR23-1A / FGSC 987;  
 RX PubMed=12712197; DOI=10.1038/nature01554;  
 RA Gajagen J.E., Calvo S.B., Borkovich K.A., Selker E.U., Read N.D.,  
 RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,  
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Qui D., Yanakiev P., Bell-Pedersen D., Nelson M.A., Braun E.L.,  
 RA Werner-Washburne M., Seltzer M.K., Kinsley J.A., Kinsley J.A.,  
 RA Zelter A., Schultze U., Koche G.O., Jedd G., Mewes H.-W., Steben C.,  
 RA Marcotte E., Greenberg D., Roy A., Foley K., Naylor J.,  
 RA Strange-Thomann N., Barrett R., Gnerre S., Kamal M., Kamysaselis M.,  
 RA Meucci E., Bielke C., Rudd S., Frishman D., Kryzofcova S.,  
 RA Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S., Cogoni C.,  
 RA Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,  
 RA Desouza C.P., Glass L., Berglund J.A., Voelker R.,  
 RA Yarden O., Plamann M., Seltzer S., Dunlap J., Radford A., Aramayo R.,  
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbols D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nisham C., Birren B.,  
 RA "The genome sequence of the filamentous fungus Neurospora crassa."  
 RL Nature 422:859-868(2003).  
 RT  
 CC -i- FUNCTION: Eliminates the production of nonsense-containing RNAs  
 (By similarity).  
 CC -i- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -i- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AL451022; CAC18314.1; -  
 CC EMBL: AABX01000272; EAA3197.1; -  
 CC InterPro: IPR005593; AAA ATPase.  
 CC InterPro: IPR001410; DEAD.  
 CC InterPro: IPR006935; Resili.  
 CC SMART: SM00382; DEXDC; 1.  
 CC SMART: SM00487; DEXDC; 1.  
 CC ATP-binding: Helicase; Hydrolyase; Hypothetical protein;  
 CC Nonsense-mediated mRNA decay; Zinc-finger.  
 CC ZN\_FING 111 139 C2H2-type (atypical) (Potential).  
 CC ZN\_FING 163 193 C4-type (Potential).  
 CC NP\_BIND 477 484 ATP (Potential).  
 CC DOMAIN 59 62 Poly-Asp.  
 CC DOMAIN 69 73 Poly-Asp.  
 CC SEQUENCE 1093 AA; 12087 MW; 8B0E4F0407ACR142 CRC64;  
 SQ  
 Query Match 92.4%; Score 73; DB 1; Length 1093;  
 Best Local Similarity 83.3%; Pred. No. 7.4e-05;  
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 IGFLEXRRINVAALTRAK 18  
 DB 835 IGFLEXRRINVAALTRAK 852  
 RESULT 9  
 YE06\_YEAST STANDARD; PRT; 1121 AA.  
 AC P32644;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Hypothetical 127.0 kDa protein in RAD24-BM1 intergenic region.  
 GN OrderedLocustNames=YE176W; ORFNames=StyG-ORF61;  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI\_TaxID=4932;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RX MEDLINE=97313264; PubMed=9169868;  
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Yelton M.A., Allen E.,  
 RA Arujo R., Aviles B., Bero A., Brennan T., Carpenter J., Chen E.,  
 RA Cherry J.M., Chung E., Duncan M., Guzman B., Hartzell G.,  
 RA Hunnicke-Smith S., Hymen R.W., Kayser A., Komp C., Lashkari D., Lew H.,  
 RA Lin D., Mosedale D., Nakahara K., Nemach A., Norgren R., Oefner P.,  
 RA Oh C., Petel F.A., Roberts D., Sehl P., Schramm S., Shogren T.,  
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;  
 RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome V."  
 RL Nature 387:78-81(1997).  
 RT  
 CC -i- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U18922; AAB64703.1; -  
 CC PIR: S30862; S30862.  
 CC Germonline: 139253.  
 CC SGD: S000000978; ECM32.  
 CC GO: GO:0005844; C:Polyhome; IDA.  
 CC GO: GO:0003678; F:DNA helicase activity; IDA.  
 CC GO: GO:0006449; P:regulation of translational termination; IMP.  
 CC InterPro: IPR001410; DEAD.  
 CC SMART: SM00487; DEXDC; 1.  
 CC ATP-binding: Helicase; Hypothetical protein.  
 CC NP\_BIND 670 677 ATP (Potential).  
 CC SEQUENCE 1121 AA; 126970 MW; 641C4AA6810282A0 CRC64;  
 SQ  
 Query Match 92.4%; Score 73; DB 1; Length 1121;  
 Best Local Similarity 83.3%; Pred. No. 7.6e-05;  
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 IGFLEXRRINVAALTRAK 18  
 DB 1065 IGFLEXRRINVAALTRAK 1082  
 RESULT 10  
 Q6BNH2 PRELIMINARY; PRT; 1124 AA.  
 AC Q6BNH2;  
 DT 25-OCT-2004 (TREMBl:rel. 28, Created)  
 DT 25-OCT-2004 (TREMBl:rel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMBl:rel. 28, Last annotation update)  
 DE Debaryomyces hanseni chromosome E of strain CBS767 of Debaryomyces  
 DE hanseni.  
 GN ORFNames=DEHA0E230349;  
 OS Debaryomyces hanseni CBS767.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.  
 OC NCBI\_TaxID=284592;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS767;  
 RG Genolevures;  
 RA Dujon B., Sherman D., Fischer G., Durieux P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla B.,  
 RA Goffard N., Franchin L., Aigle M., Anthouard V., Babour A., Barde V.,  
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boistrasse A., Boyer J., Catolico L., Confanioli F., de Daruvar A.,  
 RA Despons U., Fabre R., Fairhead C., Ferry-Dumazet H., Groppi A.,  
 RA Hantuya F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lésur I., Ma L., Muller H.,  
 RA Nicaud J.M., Nikoleki L., Oztas S., Ozier-Kalogeropoulos O.,



```

RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swemene D., Tekala F., Mesolowski-Louvel M., Meschof E., Wirth B.,
RA Zenon-Meyer M., Zivanovic I., Bolotin-Fukhar M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.,
RT "Genome evolution in yeasts."
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CS8767;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382137; CAG8524.1;
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001547; Glyco_hydro_5.
DR SMART; SM00382; AAA_1.
DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
DR ATP-binding.
SQ SEQUENCE 1124 AA; 127842 MW; D6F8B09E58614BE9 CRC64;

Query March 92.4%; Score 73; DB 2; Length 1124;
Best Local Similarity 83.3%; Pred. No. 7.7e-05;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IGFLEXRRINVALTRAK 18
Db 1067 IGFLEXRRINVALTRAK 1084

RESULT 11
ID 06FEK3 PRELIMINARY; PRT; 1125 AA.
AC 06FEK3;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Candida glabrata strain CBS138 chromosome 1 complete sequence.
GN ORFNames=CAGL01L2034g;
OS Candida glabrata CBS138.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=284593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barde V.,
RA Barney S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boistrame A., Boyer J., Catoilico L., Confantolero F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppe A.,
RA Hantreya F., Hemequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Micard J.M., Nikolaki M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swemene D., Tekala F., Mesolowski-Louvel M., Meschof E., Wirth B.,
RA Zenon-Meyer M., Zivanovic I., Bolotin-Fukhar M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.,
RT "Genome evolution in yeasts."
RL Nature 430:35-44(2004).
DR EMBL; CR380958; CAG6265.1;
SQ SEQUENCE 1125 AA; 127137 MW; 63C0428123F3CC8C CRC64;

Query March 92.4%; Score 73; DB 2; Length 1125;
Best Local Similarity 83.3%; Pred. No. 7.7e-05;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IGFLEXRRINVALTRAK 18
Db 1069 IGFLEXRRINVALTRAK 1086

```

```

RESULT 12
ID 07RKP6 PRELIMINARY; PRT; 1139 AA.
AC 07RKP6;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE SENS-related.
GN Name=PY02854;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Anguolli S.V., Suh B.B., Koolij T.W., Perteu M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmiller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabdi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Hartie M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Valdeya A.B.,
RA van Lin L.H., Jense C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.,
RT "Genome sequence and comparative analysis of the model rodent malaria
parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AAB01000797; EAA22364.1;
SQ SEQUENCE 1139 AA; 134154 MW; 51F52F81E5332E9A CRC64;

Query March 92.4%; Score 73; DB 2; Length 1139;
Best Local Similarity 83.3%; Pred. No. 7.8e-05;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IGFLEXRRINVALTRAK 18
Db 925 IGFLEXRRINVALTRAK 942

RESULT 13
ID 094387 PRELIMINARY; PRT; 1944 AA.
AC 094387;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
GN Name=SPBC29A10.10c;
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972b-;
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Baaham D., Bowman S.,
RA Brookes K., Brown D., Brown S., Chillingworth T., Churcher C.,
RA Collins M., Connor R., Cronin A., Davis P., Pelletier T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle B.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean C.,
RA Mooney P., Moutle S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch B.,

```

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Skelton J., Taylor R.G., Tivey A., Walsh S., Warren T., Whitehead S.,  
 RA Woodward J., Voiclaert G., Aert R., Robben J., Grymopier B.,  
 RA Welfens I., Vanstreels B., Bieger M., Schafer M., Muller-Auer S.,  
 RA Gabel C., Fuchs B., Dueternof A., Filiz C., Holzer E., Moestl D.,  
 RA Halbert H., Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R.,  
 RA Polt T.M., Beyer P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleaux V., Mottier S.,  
 RA Gilbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas R., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,  
 RA Daga R.R., Cruzado U., Jimenez J., Sanchez M., del Rey P., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 DR EMBL; AL034463; CAA22438.1; -.  
 DR PIR; T40065; T40065.  
 DR GeneDB; SPombe; SPBC29A10.10c; -.  
 DR InterPro; IPR000873; AMP-bind.  
 DR PROSITE; PS00455; AMP BINDING; UNKNOWN 1.  
 SQ SEQUENCE 1944 AA; 222209 MW; 12B005A34BE11C CRC64;

Query Match 92.4%; Score 73; DB 2; Length 1944;  
 Best Local Similarity 83.3%; Pred. No. 0.00013;  
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLDXRRINVAALTRAK 18  
 DB 1729 IGFLDXRRINVAALTRAK 1746

RESULT 14  
 OBJECT9 PRELIMINARY; PRT; 2743 AA.  
 AC OBJECT9;  
 DT 01-MAR-2003 (TREMblrel. 23, Created)  
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)  
 DE Hypothetical protein MAL13p1.13.  
 GN Name=MAL13p1.13;  
 OS Plasmodium falciparum (isolate 3D7).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,  
 RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,  
 RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.,  
 RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AL644509; CAD52159.1; -.  
 KM Hypothetical protein.  
 SQ SEQUENCE 2743 AA; 325169 MW; 6FAC4BCD4EF99500 CRC64;

Query Match 92.4%; Score 73; DB 2; Length 2743;  
 Best Local Similarity 83.3%; Pred. No. 0.00019;  
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IGFLDXRRINVAALTRAK 18  
 DB 2537 IGFLDXRRINVAALTRAK 2554

RESULT 15  
 Q9AVZ7 PRELIMINARY; PRT; 692 AA.  
 AC Q9AVZ7;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)  
 DE Sent protein.  
 GN Name=sen1;

OS Guillardia theta (Cryptomonas phi).  
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.  
 OX NCBI\_TaxID=55529;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20087226; PubMed=10618395; DOI=10.1073/pnas.97.1.200;  
 RA Zauner S., Fraunholz M., Wastl U., Penny S.L., Beaton M.,  
 RA Cavalier-Smith T., Maller U., Douglas S.,  
 RT "Chloroplast protein and centrosomal genes, a tRNA intron, and odd  
 telomeres in an unusually compact eukaryotic genome, the cryptomonad  
 nucleomorph.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:200-205(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21223349; PubMed=11323671; DOI=10.1038/35074092;  
 RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.,  
 RA Wu X., Reith M., Cavalier-Smith T., Maller U.,  
 RT "The highly reduced genome of an enslaved algal nucleus.";  
 RL Nature 410:1091-1096(2001).  
 DR EMBL; AJ010592; CAC27074.1; -.  
 DR PIR; E90113; E90113.  
 SQ SEQUENCE 692 AA; 81812 MW; 519717FF04C898AC CRC64;

Query Match 91.1%; Score 72; DB 2; Length 692;  
 Best Local Similarity 83.3%; Pred. No. 7.3e-05;  
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IGFLDXRRINVAALTRAK 18  
 DB 598 IGFLDXRRINVAALTRAK 615

Search completed: April 18, 2005, 08:15:49  
 Job time : 76.1533 secs